

Qualität von Datenbanken:
Eine Untersuchung der Ansätze zur Qualitätsbeurteilung anhand
ausgewählter Beispiele

Diplomarbeit
im Fach
Informationsvermittlung WB
Studiengang Wissenschaftliches Bibliothekswesen
der
Fachhochschule Stuttgart - Hochschule für Bibliotheks- und Informationswesen

Christina Fuhr, Hünfelden

Erstprüfer: Prof. Nohr
Zweitprüfer: Prof. Dr. Roos

Angefertigt in der Zeit vom 09. Juli 1999 bis 11. Oktober 1999

Hünfelden, Oktober 1999

Abstract

Die vorliegende Arbeit beschäftigt sich mit den in der Literatur beschriebenen Ansätzen zur Qualitätsbeurteilung von Online-Datenbanken und Internet-Ressourcen. Dabei werden zunächst die in Online- und Internet-Datenbanken potentiell auftretenden Fehler und Qualitätsmängel beschrieben, um dann auf bisher entwickelte Ansätze zur Qualitätsbeurteilung einzugehen und diese zu bewerten. In einem weiteren Schritt werden sodann sechs Kriterien des Ansatzes der Southern California Online User Group (SCOUG) genutzt, um die Qualität der SIGLE-Datenbank zu prüfen. Darüber hinaus wird untersucht, ob die SIGLE-Datenbank den Anforderungen an eine bibliographische Datenbank entspricht und eine ausreichende Endnutzerorientierung aufweist. Auf den Bereich der Internet-Datenbanken werden die ursprünglich zur Qualitätsbeurteilung von Internet-Ressourcen gedachten Selection criteria for quality controlled information gateways übertragen und zur Qualitätsbeurteilung der beiden CORDIS-Datenbanken *Veröffentlichungen* und *Ergebnisse* verwendet.

Schlagworte: *Datenbank, Qualität, Database Labels, SCOUG, DESIRE, SIGLE, CORDIS*

This paper examines models used to evaluate the quality of online databases and Internet resources. First, possible errors and defects that can occur in using a database are described. After that several methods to evaluate the quality of databases or Internet-Resources, described in literature, are discussed. Later six criteria from the attempt of the Southern California Online User Group (SCOUG) are used to estimate the quality of the STN-File SIGLE. In addition to that it is tested, if the SIGLE-File meets the requirements for bibliographic databases. Then the end user-orientation of the SIGLE-File is examined. After that the Selection criteria for quality controlled information gateways, originally used to evaluate the quality of Internet resources, are applied to check the quality of the CORDIS databases „*Results*“ and „*Publications*“.

Keywords: *database, quality, database labels, SCOUG, DESIRE, SIGLE, CORDIS*

Inhaltsverzeichnis

I. Gliederung

| | | |
|------|---|----|
| I. | Einleitung..... | 1 |
| 1. | Die Qualität von Datenbanken und die Konsequenzen schlechter Qualität..... | 1 |
| 2. | Vorgehensweise und Gang der Arbeit..... | 8 |
| 3. | Die Datenbanken SIGLE und CORDIS..... | 8 |
| 3.1 | Die SIGLE-Datenbank..... | 8 |
| 3.2 | Die CORDIS-Datenbanken..... | 10 |
| II. | Der Qualitätsbegriff und seine Berücksichtigung im Bibliotheks- und Informationswesen..... | 11 |
| III. | Ansätze zur Qualitätsbeurteilung und ihre Bewertung..... | 16 |
| 1. | Ansätze zur Qualitätsbeurteilung von Online-Datenbanken..... | 16 |
| 1.1 | Die Bestimmung der Vollzähligkeits- und Relevanzrate (recall und precision)..... | 16 |
| 1.2 | Die Ansätze des Komitees zur Beurteilung der Qualität von Datenbanken der Finnish Society for Information Services..... | 17 |
| 1.3 | Die Ansätze der Southern California Online User Group (SCOUG)..... | 19 |
| 1.4 | Die Ansätze des Centre for Information Quality Management (CIQM)..... | 24 |
| 2. | Ansätze zur Qualitätsbeurteilung von Internet-Dokumenten..... | 27 |
| 2.1 | Die Ansätze der WWW Virtual Library..... | 27 |
| 2.2 | Die Ansätze der Internet review sites..... | 28 |
| 2.3 | Die Ansätze des Bibliotheks- und Informationswesens..... | 29 |
| 2.4 | The Argus Clearinghouse..... | 31 |
| 2.5 | Die Ansätze des DESIRE-Projektes..... | 32 |

| | | |
|-------|---|----|
| IV. | Darstellung und Begründung der gewählten Kriterien zur Qualitätsbeurteilung..... | 35 |
| 1. | Kriterien zur Beurteilung der SIGLE-Datenbank..... | 35 |
| 2. | Kriterien zur Beurteilung der CORDIS-Datenbanken..... | 36 |
| V. | Anwendung der gewählten Kriterien auf die SIGLE- und CORDIS-Datenbanken..... | 36 |
| 1. | Beurteilung der SIGLE-Datenbank..... | 36 |
| 2. | Die Recherche in der SIGLE-Datenbank über STN Easy..... | 47 |
| 3. | Beurteilung der CORDIS-Datenbanken..... | 50 |
| VI. | Ergebnisse..... | 59 |
| VII. | Anhang..... | 61 |
| 1. | Die Ansätze des Komitees zur Qualitätsbeurteilung von Datenbanken der Finnish Society for Information Services..... | 61 |
| 2. | Die Ansätze der Southern California Online User Group (SCOUG)..... | 64 |
| 3. | Beispiel-Label des Centre for Information Quality Management..... | 71 |
| 4. | Selection criteria for quality controlled information gateways..... | 77 |
| 5. | Literatur- und Quellenverzeichnis..... | 89 |
| 5.1 | Literaturverzeichnis..... | 89 |
| 5.2 | Quellenverzeichnis..... | 94 |
| 5.3 | Verzeichnis der verwendeten Dokumentations- und Hilfsmaterialien..... | 97 |
| 5.3.1 | Materialien zu den CORDIS-Datenbanken..... | 97 |
| 5.3.2 | Materialien zur SIGLE-Datenbank..... | 98 |

II. Abkürzungsverzeichnis

| | |
|----------|---|
| CIQM | Centre for Information Quality Management |
| CORDIS | Community Research and Development Information Service |
| DESIRE | Development of a European Service for Information and Research in Education |
| GAVEL | Konsortium von europäischen Beratern in der Informationsbranche |
| EQUIP | European Quality in Information Programme |
| EUROLUG | European Online User Group |
| EUSIDIC | The European Association of Information Services |
| KMU | Kleine und mittelständische Unternehmen |
| NFAIS | National Federation of Abstracting and Information Services |
| NORDINFO | Nordic Council for Scientific Information and Research Libraries |
| SCOUG | Southern California Online User Group |
| SIGLE | System for Information on Grey Literature in Europe |
| TQM | Total Quality Management |
| UKOLUG | UK Online User Group |

III. Abbildungsverzeichnis

| | |
|--|----|
| Abbildung 1: Inkonsistenzen im Index zum Erscheinungsort..... | 3 |
| Abbildung 2: Auszug aus einer SIGLE-Recherche..... | 38 |
| Abbildung 3: SIGLE-Recherche bzgl. Konsistenz der Indexierung..... | 40 |
| Abbildung 4: Auszug aus einer SIGLE-Recherche..... | 41 |
| Abbildung 5: Verfügbare Online-Hilfen in der SIGLE-Datenbank..... | 42 |
| Abbildung 6: Recherchierbare Felder in der SIGLE-Datenbank..... | 45 |
| Abbildung 7: STN Easy-Suchmaske..... | 49 |
| Abbildung 8: Datenbank Ergebnisse..... | 52 |
| Abbildung 9: Datenbank Veröffentlichungen..... | 52 |

„With online there seems to be more of a tendency to accept answers or results, rather than to question, verify, and cross-check and many end users still believe that anything out of a computer must be good“.¹

I. Einleitung

1. Die Qualität von Datenbanken und die Konsequenzen schlechter

Qualität

In einer Zeit, in der quasi ein unbeschränkter Zugriff auf Informationen möglich ist und die Anzahl der Quellen, aus denen die Informationen entnommen werden, zudem kontinuierlich ansteigt², wird die Qualität von Informationen und der sie bereitstellenden Datenbanken zu einem wichtigen Auswahlkriterium für deren Nutzer. Insbesondere im Bereich des Internet, aber auch im Online-Bereich gilt es zu hinterfragen, welche der unzähligen Datenbanken für ein konkret zu lösendes Problem die richtige ist. Dies ist auch und vor allem deswegen wichtig, weil Online- und z.T. auch Internet-Datenbanken im Gegensatz zu CD-ROM-Datenbanken vor der eigentlichen Recherche nicht kostenfrei getestet werden können.

Die Qualität von Datenbanken, insbesondere von Online-Datenbanken, zunehmend aber auch die Qualität des Internet, werden im Bibliotheks- und Informationswesen seit längerem ausführlich diskutiert. Bereits in den fünfziger Jahren wurde mit den Werten "recall" und "precision" die Leistungsfähigkeit von Retrievalsystemen getestet.³ Neuere Ansätze entstanden Ende der achtziger und zu Beginn der neunziger Jahre. Bedauerlicherweise beschränkt sich diese neuere Diskussion jedoch überwiegend auf die USA und das europäische Ausland. An Deutschland scheint die Entwicklung bis auf wenige Stimmen, die sich Anfang der neunziger Jahre mit dem Thema "Informationsqualität" und dem Management der Qualität von Informationsdienstleistungen oder des Bibliotheks- und Informationswesens als solchem beschäftigt haben, bisher weitgehend vorübergegangen zu sein.

Die Beurteilung der Qualität von Datenbanken ist vor allem im Hinblick auf zunehmend selbständig recherchierende Endnutzer von großer Wichtigkeit. Bei Informationsfachleuten, z.B. Bibliothekaren oder Informationsvermittlern, wird in der Regel angenommen, daß sie qualitativ hochwertige Produkte erkennen; für

¹ Tenopir (1990), S. 64.

² Vgl. Armstrong (1995), S. 49.

³ Vgl. Su (1992), S. 503.

Endnutzer, wie Studenten, Wissenschaftler oder Unternehmensmanager, gilt dies jedoch nicht⁴. Besondere Probleme ergeben sich für solche Endnutzer, die nicht über ausreichende Kenntnisse zur Beurteilung einer Datenbank verfügen und daher falsche oder unvollständige Rechercheergebnisse unkritisch als korrekt akzeptieren.⁵ Zu diesem Ergebnis gelangt auch *Medawar*, wenn sie feststellt: *"In dealing with database quality, it has to be realised that the end-users are the most vulnerable of users. Not only may they be unaware of their own shortcomings, but they may not possess the skills to recognise shortcomings in the database itself, accepting as unassailable results which professional intermediaries might have questioned"*.⁶ Aus diesem Grund ist es notwendig, vor allem dem Endnutzer vorab genügend Informationen zu geben, um ihn in die Lage zu versetzen, eine zur Lösung seines konkreten Informationsproblems geeignete Datenbank auszuwählen.

Die bei der Nutzung von Datenbanken auftretenden Qualitätsmängel sind sehr zahlreich. Im folgenden sollen zunächst die in Online-Datenbanken auftretenden Qualitätsprobleme beschrieben werden, bevor auf die Problematik des Internet eingegangen wird. Dabei werden überwiegend die englischsprachigen Begriffe beibehalten, da eine Übersetzung in diesen Fällen zu Ungenauigkeiten bzw. zu einer bloßen Beschreibung des Begriffs führen würde. Da es sich aber um Fachbegriffe aus dem Bereich der Datenbankrecherche (Information-Retrieval) handelt, ist die fehlende Übersetzung insofern unschädlich.

In Online-Datenbanken treten die häufigsten Fehler und Qualitätsmängel in den Datensätzen einer Datenbank in Form von Schreibfehlern, Ungenauigkeiten, Inkonsistenzen, Indexierungsfehlern oder auch schlicht falschen Daten auf.⁷ **Schreibfehler** können in allen Bereichen und Feldern einer Datenbank auftreten. Sie können nach *O'Neill* und *Vizine-Goetz* unterteilt werden in "spelling errors", "typographical errors", "scanning errors" und "transmission errors". Ein "**spelling error**" entsteht durch falsches Buchstabieren, ein "**typographical error**" durch einen Eingabefehler. "**Scanning errors**" entstehen beim Scannen der Datensätze zugrundeliegenden Dokumente, der anschließenden Bearbeitung mit Methoden der optical character recognition (OCR) oder auch durch andere Scanning-Verfahren. "**Transmission errors**" sind Übertragungsfehler, die auftreten können, wenn die

⁴ Vgl. Armstrong (1995), S. 50.

⁵ Vgl. Armstrong (1995), S. 51.

⁶ Medawar (1995), S. 261.

⁷ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>); Armstrong (1995), S. 51; Jascó (1993 I), S. 38.

Datensätze ausschließlich telephonisch ohne entsprechende fehlerentdeckende Kommunikationsprotokolle übermittelt werden⁸.

Die sogenannten "**spelling errors**" können noch weiter unterteilt werden. *O'Neill* und *Vizine-Goetz* nennen "omission", das Auslassen von einzelnen Buchstaben; "insertion", ein Einfügen einzelner Buchstaben, "substitution", den Ersatz einzelner Buchstaben sowie "transposition", das Verdrehen von Buchstaben. Sie führen als Beispiel das Wort "great" an, bei dem als mögliche Abweichungen "gret", "greaet", "gteat" und "graet" vorkommen können.⁹ Das Auftreten und die Häufigkeiten dieser "spelling errors" sind in verschiedenen Studien weiter untersucht worden.¹⁰

Eine weitere Ausprägung der "spelling errors" entsteht nach *Mitton* dann, wenn durch den "spelling error" andere Worte entstehen, die ebenfalls einen Sinn ergeben.¹¹ Er identifiziert außerdem **Worttrennungsfehler**, bei denen Worte getrennt werden, die zusammengeschrieben werden müssten, und solche, die getrennt werden müssten, zusammengeschrieben werden.¹² Der Begriff „Information Retrieval“ würde also z.B. mit „Informationretrieval“ wiedergegeben oder der Begriff „Datenbankrecherche“ als „Datenbank-Recherche“.

Ungenauigkeiten und **Inkonsistenzen** finden sich als Folge von fehlenden Normdateien oder fehlender Standardisierung häufig in den Indizes zu Erscheinungsort und Verlag.¹³ Dies soll die nachfolgende Abbildung verdeutlichen.

| | |
|---------------------------|------------------------------------|
| Lond | London Collins in association with |
| Londdon | London Colney |
| Londen sic | London England |
| Londensic | London?p |
| Londin | LondonALBSU |
| Londino | Londonb |
| Londn | Londonb Arrow |
| Londo | LondonbBloomsbury |
| Londob | LondonbRoutledge |
| Londoin | LondonbGollancz |
| London | * |
| London 6b Severn House | LondonCalder |
| London Batsford | * |
| London Bulfinch Press... | * |
| London Business School... | LondonPenguin Books |
| London c1985 | LondonR |
| London Collier | Londons |
| London Collins | Caribbean LondonSbScolastic Press |

Abbildung 1: Inkonsistenzen im Index zum Erscheinungsort¹⁴

⁸ Vgl. *O'Neill / Vizine-Goetz* (1988), S. 132-133.

⁹ Vgl. *O'Neill / Vizine-Goetz* (1988), S. 133 in Anlehnung an *Damerau* (1964), S. 171.

¹⁰ Vgl. z.B. *Mitton* (1987), S. 495-505.

¹¹ Vgl. *Mitton* (1987), S. 496.

¹² Vgl. *Mitton* (1987), S. 498.

¹³ Vgl. *Armstrong* (1995), S. 52.

Weitere Ungenauigkeiten treten in den bibliographischen Angaben im Quellenfeld auf. Diese Fehler sind jedoch meist schon im Ursprungsdokument vorhanden¹⁵ und entstehen nicht beim Erfassen der Daten für eine Datenbank.

Inkonsistenzen finden sich zudem nicht nur in den Datensätzen einer Datenbank, auch die zwischen verschiedenen Datenbanken ständig wechselnden Feldbezeichnungen stellen Inkonsistenzen dar. In einer Datenbank bezeichnet eine bestimmte Feldabkürzung das Firmenfeld ("Company"), in einer anderen ist damit das Feld für die Zeitschriftenkurzform ("Codex") gemeint. Die unterschiedliche Suchbarkeit von bestimmten Feldern in verschiedenen Datenbanken wird in diesem Zusammenhang ebenfalls angeführt.¹⁶

Ein weiteres Problem tritt im Bereich von **"Coverage and scope"** auf. Der Begriff "Coverage and scope" bezeichnet den Themenbereich oder das Fachgebiet der Datenbank sowie den geographischen Raum, aus dem die ausgewerteten Quellen stammen. Nutzer, insbesondere Endnutzer, finden in der Datenbank häufig nicht die Quellen vor, die sie in einem bestimmten Fachgebiet erwarten. Die Quellen werden in unterschiedlicher Breite ausgewertet, Zeitschriften z.B. nur selektiv oder cover-to-cover. Außerdem kommt es vor, daß die für das Fachgebiet wichtigste Quelle überhaupt nicht ausgewertet wird. Die Auswertungs- und Indexierungspraxis der Datenbankproduzenten ändert sich zudem im Laufe der Zeit, ohne daß vorgenommene Änderungen auch retrospektiv erfaßt werden.¹⁷ Die aktuelle Auswertungspraxis für eine bestimmte Datenbank zu ermitteln, ist für einen Nutzer nur sehr schwer, z.B. durch Lesen von langen und umständlichen Handbüchern oder Bluesheets, möglich. Häufig erläutern die Produzenten einer Datenbank ihre Auswertungspraxis nicht bzw. versprechen in Werbeprospekten Inhalte, die sich bei näherem Hinsehen gar nicht oder nicht in der versprochenen Breite finden lassen.

Einen zusätzlichen potentiellen Problembereich stellt die **Aktualität** einer Datenbank dar.¹⁸ Die Aktualität einer Datenbank kann als *"Zeitraum, der zwischen dem Erscheinen der Primärinformation bis zur Verfügbarkeit der Sekundärinformation in der Datenbank vergeht"*¹⁹ definiert werden. In diesem Zusammenhang wird insbesondere auch das Problem der unterschiedlichen Aktualität von Datenbank und

¹⁴ Abbildung nach Armstrong (1995), S. 52 leicht modifiziert.

¹⁵ Vgl. Jascó (1997), S. 252 sowie Mintz (1990), S. 15.

¹⁶ Vgl. Basch (1992), S. 23.

¹⁷ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

¹⁸ Vgl. Armstrong (1995), S. 52.

¹⁹ Norpoth / Nüßlein (1995), S. 208.

Quellen genannt. Die einzelnen einer Datenbank zugrundeliegenden Quellen können eine sehr unterschiedliche Aktualität aufweisen.²⁰

Qualitätsmängel können auch durch das Fehlen wichtiger Felder („**errors of omission**“) entstehen, die bei der Nutzung einer Datenbank unbedingt erwartet werden, so z.B. Erscheinungsjahr, Dokumententyp, oder Sprachangabe.

Probleme ergeben sich ebenfalls dann, wenn einzelne Felder dazu verwendet werden, alle Daten zu erfassen, die in keines der anderen Felder eingetragen werden können.²¹

Als weiterer Qualitätsmangel ist die Tatsache zu nennen, daß sowohl professionelle Rechercheure als auch Endnutzer noch immer **Retrievalssprachen** erlernen müssen, um eine Datenbank nutzen zu können.²² Retrievalssprachen erschweren die Bedienbarkeit einer Datenbank und damit den Zugang zu den Daten. Sie entsprechen damit nicht dem erforderlichen "**ease of use**".²³

Weiterhin werden die **Ausgabemöglichkeiten** vieler Datenbanken beanstandet. Ein Ranking nach der Relevanz erzielter Suchergebnisse ist meist nicht möglich, und auch die Standardausgabeformate lassen zu wünschen übrig.²⁴

Mangelhaft kann auch die **Dokumentation** einer Datenbank sein; Online-Hilfe und gedruckte Hilfsmittel sollten den gleichen Stand aufweisen.²⁵ Basch weist darauf hin, daß die Online-Hilfe kontextsensitiv sein sollte. **Fehlermeldungen** sollten in gesprochener und nicht in Programmiersprache abgefaßt sein.²⁶

Eine qualitativ hochwertige Datenbank weist außerdem ein vernünftiges Preis-Leistungsverhältnis auf ("**Value-to-Cost Ratio**").²⁷ In diesem Zusammenhang ist insbesondere die generelle Preisgestaltung von Hosts zu kritisieren, die den Preis für eine Online-Recherche nach der Anschlußzeit bestimmen.²⁸ Dies ist besonders für ungeübte Nutzer ein großer Nachteil.

Fehler und Qualitätsmängel in Datenbanken haben sowohl Konsequenzen für den Nutzer, als auch für die Produzenten und Anbieter einer Datenbank. Die genannten Fehler können dazu führen, daß in der Datenbank verzeichnete relevante Quellen

²⁰ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

²¹ Vgl. Jascó (1993 I), S. 38.

²² Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

²³ Vgl. Armstrong (1995), S. 52.

²⁴ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>); Armstrong (1995), S. 52.

²⁵ Vgl. Armstrong (1995), S. 52.

²⁶ Vgl. Basch (1992), S. 24-25.

²⁷ Vgl. Armstrong (1995), S. 52.

²⁸ Vgl. Basch (1992), S. 23.

nicht ermittelt werden können. Hierdurch bleiben entscheidende Informationen möglicherweise unberücksichtigt.²⁹ Dies z.B. im Falle einer Patentrecherche zu großen finanziellen Schäden führen und die Wettbewerbsfähigkeit eines Unternehmens beeinträchtigen. Insbesondere dann, wenn Gelder in Forschungsprojekte investiert werden, für die bereits ein Patent erteilt wurde. Fehler in Faktendatenbanken, z.B. bezüglich der Zusammensetzung von chemischen Substanzen oder der notwendigen Dosierung von Medikamenten könnten verheerende Folgen haben. In bibliographischen Datenbanken erschweren oder verhindern diese Fehler das Auffinden der in den Datenbanken zitierten Quellen.³⁰

Neben dem möglichen Verlust von Informationen können die Qualitätsmängel oder Fehler aber auch dazu führen, daß zu viele Datensätze angezeigt werden. Das Centre for Information Quality Management (CIQM) ermittelte in einer Studie, daß nach einer Datenbankrecherche häufig irrelevante und ungewollte Daten aussortiert werden müssen, was die für die Suche notwendige Zeit verlängert. Ausgabeprobleme führen zudem dazu, daß relevante Ergebnisse aufgrund von Formatierungsproblemen für irrelevant erachtet werden. Im Extremfall bewirkt dies, daß Recherchen wiederholt werden müssen, wodurch zusätzliche Kosten verursacht werden.³¹

Auch eine rechtliche Verantwortlichkeit für fehlerhafte Informationsprodukte wird zunehmend diskutiert. In der Europäischen Union gibt es Ansätze, die Grundsätze der Produkthaftung auch auf die Informationsindustrie auszuweiten.³² Hinzu kommt, daß Fehler und Qualitätsmängel hohe Kosten durch erforderliche Nachbesserungen und mögliche Kundenverluste verursachen können. Für die Informationsindustrie stellen qualitativ hochwertige Informationsprodukte, z.B. Datenbanken, bedingt durch die für Nutzer auf dem Markt vorhandenen großen Auswahlmöglichkeiten, zudem einen Wettbewerbsfaktor dar.³³

Bei Benutzung des Internet entstehen insbesondere bei der Suche nach wissenschaftlichen Informationen andersartige Probleme. Mit der Entstehung des World Wide Web wurden der Zugang zum und das Publizieren im Internet vereinfacht.³⁴ Die Anzahl der Nutzer und der Seiten steigt kontinuierlich an, wodurch das Hauptproblem des Internet, das sogenannte "Information overload", entsteht.³⁵

²⁹ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

³⁰ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

³¹ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

³² Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

³³ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>).

³⁴ Vgl. Kirk (1999) <http://milton.mse.jhu.edu:8001/research/education/net.html>.

³⁵ Stellvertretend für zahlreiche Quellen: IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>); Chung / Lee 1998, S. 146.

Durch die Fülle von Informationen ist es schwierig und zeitaufwendig, relevante und qualitativ hochwertige Informationen zu finden.³⁶ Dies liegt unter anderem auch darin begründet, daß praktisch jeder im World Wide Web publizieren kann. Häufig ist es nicht möglich festzustellen, welcher Autor, welche Organisation oder welche Gruppe hinter einem bestimmten Dokument steht (Authority). Ist dagegen ein Urheber festzustellen, sind häufig seine Qualifikationen und Ziele nicht zu ermitteln (Objectivity). Bei vielen Internet-Dokumenten ist zudem aufgrund fehlender Daten die Aktualität nur schwer oder gar nicht feststellbar. Sind Daten vorhanden, können sie verschiedene Bedeutung haben: Datum der ersten Erstellung des Textes, Tag der zur Verfügungstellung im Netz, Datum der letzten Überprüfung. Nicht selten ist auch die Breite und Tiefe, mit der ein dargestelltes Thema behandelt wird, nur schwer zu fassen (Coverage).³⁷

Auch im Bereich des Internet ist die Genauigkeit der Daten ein wichtiger Aspekt, die jederzeit verifizierbar sein muß (Accuracy or verifiability)³⁸. Selbstverständlich sollten die Daten auch fehlerfrei und die enthaltene Information vollständig sein.³⁹

Die soeben beschriebenen allgemeinen Qualitätsprobleme im Internet können nach Meinung der Verfasserin auch auf Internet-Datenbanken im speziellen übertragen werden. Bei vielen dieser Verzeichnisse im Netz ist ebenfalls nicht erkennbar, durch wen sie erstellt wurden oder wie aktuell sie sind. Noch schwerwiegender ist in diesem Zusammenhang aber, daß häufig nicht feststellbar ist, welche Auswahlkriterien den Datenbanken zugrunde liegen. Viele Internet-Datenbanken werden direkt durch ihre Produzenten ins Internet gestellt, so daß typischerweise die von Online-Datenbanken bekannten Hintergrundinformationen und Dokumentationen fehlen.⁴⁰

Aufgrund der genannten Probleme kann die Qualität der im Internet gefundenen Informationen stark variieren. Eine Qualitätsbeurteilung ist daher unbedingt notwendig, aber aufgrund von ,im Gegensatz zu Online-Datenbanken, häufig nicht vorhandenen strukturierten Feldern⁴¹ ungleich schwieriger. Derzeit liegt die Qualitätsbeurteilung vielfach noch in der Verantwortung der Nutzer, die dabei allerdings auf zahlreiche bereits entwickelte Ansätze zurückgreifen können.⁴² Die diesen Modellen zugrunde gelegten Kriterien zur Beurteilung sind jedoch so

³⁶ Vgl. Cooke et. al. (1996), S. 34 ; Ochsner / Thomas (1996), S. 478.

³⁷ Vgl. Alexander / Tate (<http://www.widener.edu/?withers/webeval.htm>).

³⁸ Kirk (1999) <http://milton.mse.jhu.edu:8001/reseach/education/net.html>.

³⁹ Grassian (1999) <http://www.library.ucla.edu/libraries/college/instruct/web/critical.htm>.

⁴⁰ Vgl. Jascó (1997), S. 256.

⁴¹ Vgl. Jascó (1997), S. 256.

⁴² Vgl. Armstrong (1997) http://www.i-a-l.co.uk/ciqm/db_qual/2_1_4.htm.

verschieden, daß dringend ein einheitlicher Ansatz zur Beurteilung verwendet werden sollte.

2. Vorgehensweise und Gang der Arbeit

Die vorliegende Arbeit setzt sich mit verschiedenen bereits entwickelten Ansätzen zur Qualitätsbeurteilung von Datenbanken auseinander. Dazu wird zunächst die relevante Literatur zu den in Online-Datenbanken und im Internet auftretenden Qualitätsproblemen ausgewertet.⁴³ In einem zweiten Schritt soll auf den Qualitätsbegriff und seine Berücksichtigung im Bibliotheks- und Informationswesen eingegangen werden. Danach werden die in der Literatur vorhandenen Ansätze zur Qualitätsbeurteilung beschrieben und bewertet. Im Bereich des Internet ist es jedoch aufgrund der Vielzahl der verschiedenen Modelle nicht möglich, eine abschließende Übersicht zu geben. Die Arbeit beschränkt sich daher auf einige signifikante Beispiele, an denen die Grundtendenzen deutlich gemacht werden können. Im Anschluß an die Beschreibung der vorhandenen Ansätze werden konkrete Kriterien ausgewählt, um diese auf drei Beispiele, eine Online- und zwei Internet-Datenbanken, zu übertragen.

In allen Abschnitten der Arbeit, sowohl bei der Beschreibung der Qualitätsmängel und Fehler als auch bei der Darstellung der vorhandenen Ansätze zur Qualitätsbeurteilung, werden dabei in der Regel die englischsprachigen Begriffe beibehalten.⁴⁴

3. Die Datenbanken SIGLE und CORDIS

Als Beispiele zur Anwendung der Kriterien wurden die Datenbanken SIGLE und CORDIS gewählt, da es sich in beiden Fällen um charakteristische Beispiele handelt, die in ihrer inhaltlichen Ausprägung einzigartig und aufgrund dessen von besonderer Bedeutung sind.

3.1 Die SIGLE-Datenbank

SIGLE (System for Information on Grey Literature in Europe) ist eine bibliographische Datenbank, die über den Host STN abfragbar ist. Die Datenbank verzeichnet

⁴³ Vgl. Abschnitt 1.1

europäische nicht-konventionelle, sogenannte graue Literatur. Sie umfaßt die Bereiche der theoretischen und angewandten Wissenschaften sowie der Technik. Die Berichtszeit der Datenbank beginnt 1980, ab 1984 sind auch die Wirtschafts- und Sozialwissenschaften verzeichnet. SIGLE weist unter anderem folgende Dokumenttypen nach:

?? Forschungsberichte

?? Kongressberichte

?? Dissertationen

?? sonstige nicht-konventionelle Literatur⁴⁵.

In den Bearbeitungszeitraum der vorliegenden Arbeit fiel die Integration der STN-Datenbank FTN (Forschungsberichte aus Technik und Naturwissenschaften) in die SIGLE-Datenbank.

Die gesamte in SIGLE verzeichnete Literatur ist in der Regel nicht über den Buchhandel zu beziehen.⁴⁶ Damit ist eine für graue Literatur generell sehr schwierige Nachweissituation verbunden. Durch SIGLE soll die Nachweissituation europäischer grauer Literatur verbessert werden. Da sie ist auf europäischer Ebene den einzigen Nachweis dieser Art darstellt, ist die Qualität der Datenbank von herausragender Bedeutung.

Zur Qualitätsbeurteilung werden im Rahmen dieser Arbeit Recherchen in der SIGLE-Datenbank durchgeführt. Zugriffsmöglichkeiten bestehen für die gesamte Bearbeitungsdauer. Zur Unterstützung der Recherchen stehen die durch den Anbieter der Datenbank herausgegebenen „Summary Sheets“ sowie die sonstigen Dokumentationsmaterialien zur Verfügung.

Im Rahmen der Qualitätsbeurteilung geht die vorliegende Arbeit zu Vergleichszwecken auch kurz auf eine zweite Angebotsform der Datenbank ein, die SIGLE-Recherche über STN Easy.

⁴⁴ Zur Begründung vgl. Abschnitt 1.1.

⁴⁵ Vgl. SIGLE Summary Sheet (1999) unter <http://info.cas.org/ONLINE/DBSS/sigless.html>.

⁴⁶ Vgl. SIGLE Summary Sheet (1999) <http://info.cas.org/ONLINE/DBSS/sigless.html>.

3.2 Die CORDIS-Datenbanken

CORDIS (Community Research and Development Information Service) ist ein Datenbanksystem bzw. *"der Informationsdienst für von der Gemeinschaft geförderte Forschung und Entwicklung"*.⁴⁷ CORDIS möchte nach eigener Aussage insbesondere kleinen und mittelständischen Unternehmen die Möglichkeit geben, sich über Forschungs- und Entwicklungsaktivitäten in der Europäischen Union auf dem laufenden zu halten.⁴⁸

CORDIS besteht aus zehn verschiedenen Datenbanken:

?? Nachrichten

Verzeichnis tagesaktueller Informationen zu Forschungs- und Entwicklungsaktivitäten der Europäischen Gemeinschaften, z.B. Aufrufe zu Vorschlägen, Ausschreibungen, Veranstaltungen;

?? Programme

Informationen zu EU-geförderten Forschungsprogrammen und EU-Programmen, die sich auf Forschung und Entwicklung beziehen;

?? Projekte

Informationen zu einzelnen im Rahmen der EU-Programme ausgeführten Forschungsprojekten;

?? Partner

Verzeichnis zur Unterstützung bei der Suche nach Projektpartnern;

?? Ergebnisse

Verzeichnis zur Ermittlung von Forschungsergebnissen und Prototypen;

?? Kontakte

Liste von offiziellen Beratungsstellen, die bei EU-geförderten Forschungs- und Entwicklungsprojekten informieren und beraten;

?? Veröffentlichungen

Enthält *„Studien, Berichte und wissenschaftliche Beiträge zur EU-Forschung“*;

?? Akronyme

Akronymenverzeichnis zu Begriffen aus dem Bereich der Forschung und Entwicklung der EU;

⁴⁷ CORDIS-Einführung (1999) http://www.cordis.lu/de/src/s_016_de.htm.

?? KOM-Dokumente

Verzeichnis von Dokumenten im Zusammenhang mit den Rechtsetzungs- und Entscheidungsverfahren der Europäischen Union sowie die

?? Dokumentenbibliothek

Verzeichnis offizieller Dokumente, z.B. Aufrufe zu Vorschlägen, Arbeitsprogramme, Informationspakete, Teilnahmeleitlinien.⁴⁹

In ihrer inhaltlichen Ausprägung sind die CORDIS-Datenbanken ebenfalls einzigartig. Sie sind der einzige umfassende Nachweis zu Forschungs- und Entwicklungsaktivitäten der Europäischen Union. Die Qualität der Datenbanken ist deshalb ebenfalls sehr wichtig.

Vorliegende Arbeit wird sich bei der Anwendung der Qualitätskriterien⁵⁰ auf die Datenbanken *Ergebnisse* und *Veröffentlichungen* beschränken. Beide Datenbanken sind nach Meinung der Verfasserin für die Zielgruppe des CORDIS-Systems besonders interessant. Durch die Verzeichnung der Forschungsergebnisse (Datenbank *Ergebnisse*) können sich Unternehmen über neue Entwicklungen ihrer Branche informieren. Die Datenbank *Veröffentlichungen*, verzeichnet die Publikationen, mit deren Hilfe sich die Zielgruppe über die Forschungs- und Entwicklungsaktivitäten in der Europäischen Union informieren kann. Zudem bietet sie mit ca. 80.000 Datensätzen⁵¹ eine breite Ausgangsbasis zur Qualitätsbeurteilung. Zur Qualitätsbeurteilung werden auch in den CORDIS-Datenbanken Datenbankrecherchen durchgeführt. Als Hilfsmittel dienen dabei die durch CORDIS im Internet veröffentlichten Materialien.⁵²

II. Der Qualitätsbegriff und seine Berücksichtigung im Bibliotheks- und Informationswesen

Bei der Beschäftigung mit der Qualität von Datenbanken und den Kriterien zu ihrer Beurteilung ist es zunächst notwendig, sich auch mit dem Qualitätsbegriff und seiner Berücksichtigung im Informationssektor auseinanderzusetzen. Das Problem bei der

⁴⁸ CORDIS-Einführung (1999) http://www.cordis.lu/de/src/s_016_de.htm.

⁴⁹ Vgl. CORDIS-Einführung (1999) http://www.cordis.lu/de/src/s_016_de.htm.

⁵⁰ Vgl. Kapitel V.

⁵¹ Die Anzahl der Datensätze wurde durch eine Recherche in der Datenbank ermittelt.

⁵² Die Datenbanken und Hilfsmaterialien sind über die Homepage von CORDIS unter <http://www.cordis.lu> zugänglich.

Bestimmung des Qualitätsbegriffs besteht darin, daß Qualität schwer zu definieren oder zu messen ist. Dennoch wird häufig, so auch im Bibliotheks- und Informationswesen, behauptet, daß man sie erkenne, wenn sie vorliege.⁵³ Die Schwierigkeiten in der Bestimmung sind in einem subjektiven Ansatz begründet, den jede Qualitätsdefinition zwangsläufig beinhaltet.

Damit Qualität gemessen werden kann, z.B. bei der Beurteilung von Datenbanken, müssen jedoch objektive Ansätze vorliegen bzw. die Qualität vergleichbar oder, wie *Pepels* sagt, operationalisierbar gemacht werden.⁵⁴

Bei einer Definition des Qualitätsbegriffs können fünf verschiedene Vorgehensweisen unterschieden werden: der transzendente, der produktbezogene, der anwenderbezogene, der prozeßbezogene und der wertorientierte Qualitätsansatz.⁵⁵

Die **transzendente** Sichtweise stellt einen absoluten Qualitätsansatz dar. Qualität wird hierbei bestimmt durch die Exzellenz oder Vortrefflichkeit einer Leistung. Diese Vortrefflichkeit näher zu bestimmen oder zu messen, ist jedoch sehr schwierig.⁵⁶ Es handelt sich somit um einen subjektiven Ansatz, der zu einer Qualitätsmessung nicht ohne weitere Konkretisierungen herangezogen werden kann.

Beim **produktbezogenen** Ansatz wird Qualität durch konkrete Eigenschaften eines Produkts bestimmt, wodurch mögliche Qualitätsabweichungen objektiv und genau meßbar sind.⁵⁷

Bei der **prozeßbezogenen** Sichtweise entsteht Qualität durch die Einhaltung von Spezifikationen. Qualitativ hochwertige Produkte liegen vor, wenn die definierten Anforderungen erfüllt wurden.⁵⁸

Beim **wertorientierten** Qualitätsansatz wird Qualität als durch Kosten und Leistung bestimmt betrachtet. Qualität liegt vor, wenn eine bestimmte Leistung zu einem angemessenen Preis erbracht wird.⁵⁹

Die **anwenderbezogene** Sichtweise betrachtet Qualität aus der Sicht eines Nutzers des Produkts. Die Produkte, welche die Bedürfnisse des Kunden am besten erfüllen,

⁵³ Vgl. Johannsen (1995), S. 319.

⁵⁴ Vgl. Pepels (1996), S. 39.

⁵⁵ Vgl. Garvin (1988), S. 39-46.

⁵⁶ Vgl. Stroetmann (1993), S. 12.

⁵⁷ Vgl. Stroetmann (1993), S. 12.

⁵⁸ Vgl. Pepels (1996), S. 39.

⁵⁹ Vgl. Pepels (1996), S. 39.

weisen die höchste Qualität auf⁶⁰. Dieser subjektive Ansatz wird auch als "fitness for purpose"⁶¹ bezeichnet. Dabei geht es nicht nur darum,

⁶⁰ Vgl. Pepels (1996), S. 39.

⁶¹ Juran zitiert nach Johannsen (1995), S. 322.

offensichtliche Erwartungen zu erfüllen, sondern diese durch das Erkennen weiterer unausgesprochener Bedürfnisse zu übertreffen.⁶²

Die Industrie beschäftigt sich schon seit längerer Zeit mit dem Qualitätsbegriff und definiert Qualität in einer Kombination verschiedener Sichtweisen z.B. als *"establishing the characteristics of a product or service so that the requirements of the customer are satisfied at an acceptable cost"*.⁶³ Eine weitere Definition findet sich in der ISO-Vorschrift 8402, in der Qualität mit *"the totality of characteristics of an entity that bears on its ability to satisfy stated and implemented needs"*⁶⁴ festgelegt wird. Beide Definitionen beinhalten Ansätze, welche die Qualität meßbar machen. Qualität bestimmt sich durch den Grad der Übereinstimmung mit den geforderten Eigenschaften.

In der Informationsbranche begann die Beschäftigung mit dem Qualitätsbegriff erst relativ spät. Dennoch ist die Qualitätsdiskussion auch hier inzwischen weit verbreitet.⁶⁵ Dabei bedient sich die Informationsbranche den Ansätzen der Industrie. Ähnlich wie in der Industrie, existieren im Informationssektor verschiedene Sichtweisen und Definitionen des Qualitätsbegriffs. *Aitchison* z.B. beschreibt Qualität als *"degree of excellence"*, der eine Idealvorstellung darstellt und in fünf verschiedene Unterpunkte (*"desiderata"*) weiter aufgespalten wird. Er legt „absolute accuracy“ (absolute Genauigkeit), „complete reliability“ (Verlässlichkeit), „unvarying consistency“ (Konsistenz), „total comprehensiveness“ (Vollständigkeit) und „maximum timeliness“ (größtmögliche Aktualität) als Qualitätsmerkmale fest. Die Qualität bestimmt sich auch hier durch den Grad der Übereinstimmung mit dem Ideal.⁶⁶ Für *Keiser* steht fest, daß der Kunde die Qualität festlegt.⁶⁷ Zum gleichen Ergebnis kommt *Basch*⁶⁸; aber auch *Hobohm*⁶⁹ diskutiert die kundenorientierte Sichtweise des Qualitätsbegriffs, die damit den Schwerpunkt darzustellen scheint.

Innerhalb der Qualitätsbewegung der Informationsbranche können grundsätzlich zwei Strömungen unterschieden werden. Die erste betrifft die Einführung des Qualitätsmanagements in Einrichtungen des Informationssektors, die zweite

⁶² Vgl. Johannsen (1995), S. 323.

⁶³ Swindells (1995) <http://www.echo.lu/qa/en/qa2.html>

⁶⁴ Norm EN ISO 8402 (1995), S. 9.

⁶⁵ Vgl. Hobohm (1998) <http://www.fh-potsdam.de/~hobohm/gib.pdf>

⁶⁶ Vgl. Aitchison (1988), S. 50f..

⁶⁷ Vgl. Keiser (1993), S. 47.

⁶⁸ Vgl. Basch (1995), S. 10.

⁶⁹ Vgl. Hobohm (1998) <http://www.fh-potsdam.de/~hobohm/gib.pdf>

beschäftigt sich mit der Qualität von Datenbanken⁷⁰ - dem Thema der hier vorliegenden Arbeit.

Die Einführung von Maßnahmen des Qualitätsmanagements in Einrichtungen der Informationsbranche begann in den achtziger Jahren.⁷¹ Dabei waren vor allem die USA und die Länder Skandinaviens führend. Es entstanden verschiedene Initiativen und Aktionsprogramme, z.B. ein Forschungsprojekt von NORDINFO (Nordic Council for Scientific Information and Research Libraries) zur Untersuchung der Anwendung von Maßnahmen des TQM (Total Quality Managements) in Einrichtungen des Bibliothekswesens oder die Arbeit des EQUIP-Konsortiums (European Quality in Information Programme), bei dem EUSIDIC (The European Association of Information Services), GAVEL (ein Konsortium von europäischen Beratern im Bereich der Informationsbranche) und EUROLUG, die European Online User Group zusammenarbeiten. Die Arbeit des EQUIP-Konsortiums führte zur Planung eines Qualitätsmanagementprogramms für den Informationssektor, das der Generaldirektion XIII der Europäischen Kommission vorgelegt und in der Zeit von Januar 1993 bis April 1994 durchgeführt wurde.⁷² 1994 und 1995 richtete die Generaldirektion XIII dann jeweils Konferenzen aus, bei denen die Qualität von Informationsprodukten und -diensten diskutiert wurde. Weitere Maßnahmen der Kommission wurden z.B. mit dem INFO2000-Programm verabschiedet⁷³.

Mit dem zweiten Strang der Qualitätsbewegung, der Untersuchung der Qualität von Datenbanken, beschäftigen sich verschiedene Gruppen innerhalb des Bibliotheks- und Informationswesens. Dazu gehören das Komitee zur Beurteilung der Qualität von Datenbanken der Finnish Society for Information Services, die Arbeit der Southern California Online User Group (SCOUG) und die Zusammenarbeit der UK Library Association mit der UK Online User Group (UKOLUG), die zur Gründung des Centre for Information Quality Management (CIQM) führte. Die Arbeit dieser Gruppen bzw. ihre Ansätze zur Qualitätsbeurteilung dürften von zwei industriellen Modellen zur Qualitätsmessung beeinflusst sein: dem Modell zur Messung der Produktqualität und dem Modell zur Feststellung der Servicequalität.

Das Modell zur Messung der Produktqualität wurde von *Marchand* auf das Bibliotheks- und Informationswesen übertragen. *Marchands* Ausgangspunkt bilden

⁷⁰ Vgl. Gilchrist (1995), S. 251.

⁷¹ Vgl. Johannsen (1995), S. 335.

⁷² Vgl. Wilson (1998), S. 345.

⁷³ Vgl. IMO working paper 95/4 <http://www2.echo.lu/impact/imo/9504.html>.

hierbei die zuvor von *Garvin* entwickelten Grundsätze (dimensions) zur Bestimmung der Produktqualität. Zur Messung der Informationsqualität unterscheidet

Marchand: „actual value“, „features“, „reliability“, „meaning over time“, „relevance“, „validity“, „aesthetics“ und „percieved value“. ⁷⁴

Die erste Dimension, der **actual value**, verkörpert den Wert, den ein Informationsprodukt oder eine Informationsdienstleistung für den Nutzer hat. Dieser Wert ist an die Eigenschaften des Produkts gebunden, die seinen Nutzen begründen. **Features** stellen die Grundeigenschaften eines Produkts oder einer Dienstleistung dar. **Reliability**, die dritte Ebene zur Beurteilung, bestimmt den Grad der Verlässlichkeit des Informationsproduktes oder -services. **Meaning over time** bezeichnet die Dauerhaftigkeit eines Produktes oder einer Dienstleistung. Die fünfte Ebene der Qualitätsbeurteilung ist durch die **Relevanz** bestimmt, die den Grad der Übereinstimmung mit den Qualitätserwartungen des Nutzers angibt. Die sechste Ebene bewertet die **Gültigkeit** der Information. Die siebte Dimension stellen nach Marchand die **aesthetics** dar. Diese sind sehr subjektiv durch die Art der Präsentation und Lieferung des Produkts, also das äußere Erscheinungsbild bestimmt. Die achte Ebene verkörpert der **percieved value**. Da ein Nutzer selten alle Eigenschaften eines Produktes oder einer Dienstleistung kennt, ist er auf indirekten Qualitätsmessungen in Form von Vergleichen angewiesen. Dabei kann dann die Reputation eines Produktes oder einer Dienstleistung eine entscheidende Rolle spielen.

Ein zweites aus der Industrie stammendes Modell zur Qualitätsmessung, welches im Bibliotheks- und Informationswesen diskutiert wurde, ist das von *Parasuraman*, *Zeithaml* und *Berry* entwickelte SERVQUAL-Modell zur Bestimmung der Servicequalität. Es umfaßt fünf Ebenen: tangibles, reliability, responsiveness, assurance und empathy. ⁷⁵

Der Begriff **Tangibles** beschreibt das Gesamterscheinungsbild der Dienstleistung. Die Kategorie **Reliability** mißt die Einhaltung des Leistungsversprechens, der Begriff **responsiveness** bezeichnet den Leistungswillen des Anbieters der Dienstleistung. Mit **assurance** wird die Leistungskompetenz gemessen und **empathy** gibt die Bereitschaft wieder, sich um die individuellen Kundenwünsche zu kümmern. ⁷⁶ Insbesondere das EQUIP-Konsortiums hat sich im Rahmen eines Projektes intensiv mit den Ansätzen des SERVQUAL-Modells beschäftigt. ⁷⁷ Ziel des Projektes war es, Methoden zu entwickeln, mit denen die Zufriedenheit von Datenbanknutzern

⁷⁴ Vgl. Marchand (1990), S. 10-13.

⁷⁵ Vgl. Parasuraman et.al. (1988), S. 6.

⁷⁶ Vgl. Haller (1993), S. 23.

⁷⁷ Vgl. Wilson (1998), S. 346.

gemessen werden sollte. Dabei kamen die Projektteilnehmer zu dem Ergebnis, daß eine direkte Übertragung des SERVQUAL-Modells auf den Datenbankbereich nicht möglich ist. Übernommen wurde aber die von Parasuraman, Zeithaml und Berry entwickelte Begrifflichkeit.⁷⁸

Im folgenden sollen die von der SCOUG, dem Komitee zur Beurteilung der Qualität von Datenbanken der Finish Society for Information Services und dem CIQM entwickelten Ansätze zur Beurteilung der Qualität von Datenbanken dargestellt werden.

III. Ansätze zur Qualitätsbeurteilung von Datenbanken und ihre Bewertung

1. Ansätze zur Beurteilung von Online-Datenbanken

1.1 Die Bestimmung der Relevanz- und Vollzähligkeitsrate (recall und precision)

Erste Ansätze zur Beurteilung der Qualität von Datenbanken gab es bereits in den fünfziger Jahren. Ein typisches Beispiel ist die Bestimmung von "recall" und "precision" zur Feststellung der Leistungsfähigkeit von Retrievalsystemen.⁷⁹ **Recall** bezeichnet die Vollzähligkeitsrate, **precision** die Relevanzrate einer Datenbank. Beide Werte können berechnet werden. Die **Relevanzrate** läßt sich bestimmen, indem die Menge der bei einer Recherche ermittelten relevanten Dokumente durch die Gesamtzahl der ermittelten Dokumente dividiert wird. Die **Vollzähligkeitsrate** ergibt sich, wenn den ermittelten relevanten Dokumenten die Gesamtzahl der in der Datenbank gespeicherten relevanten Dokumente gegenübergestellt wird.⁸⁰ Zur Ermittlung der Vollzähligkeitsrate, müßte somit die Gesamtzahl der gespeicherten relevanten Dokumente bekannt sein.⁸¹ Da dies jedoch nicht möglich ist und mit diesem Ansatz zur Qualitätsbeurteilung zudem z.B. nicht die Konsistenz oder der "ease of use" einer Datenbank festgestellt werden kann, soll er im Rahmen dieser Arbeit nicht weiterverfolgt werden.

⁷⁸ Vgl. Wilson (1998), S. 352.

⁷⁹ Vgl. Su (1992), S. 503.

⁸⁰ Vgl. Gaus (1995), S. 225.

⁸¹ Vgl. Su (1992), S. 503.

1.2 Die Ansätze des Komitees zur Beurteilung der Qualität von Datenbanken der Finnish Society for Information Services

Das Komitee zur Beurteilung der Qualität von Datenbanken der Finnish Society for Information Services wurde 1988 als Arbeitsgruppe gegründet, um die Qualität finnischer Datenbanken zu studieren und zu beurteilen.⁸² Die Beschäftigung mit Online-Systemen begann in Finnland bereits in den siebziger Jahren; die ersten eigenen Datenbanken wurden 1975 entwickelt. Generell ist in Finnland eine sehr hohe Benutzungsrate sowohl von ausländischen als auch von eigenen Datenbanken zu verzeichnen.⁸³ Durch die in der finnischen Online-Industrie gering ausgeprägte Standardisierung kam es schon Ende der achtziger Jahre zur Entstehung einer Konsumentenbewegung,⁸⁴ mit der auch die Gründung des Komitees zur Beurteilung der Qualität von Datenbanken in Zusammenhang steht. Dieses Komitee entwickelte im Rahmen eines im Februar 1989 begonnenen Projektes, eine Liste mit Qualitätskriterien, die zur Beurteilung von Datenbanken genutzt werden sollte. Diese Liste entstand durch die Auswertung verschiedener Studien zum Thema Datenbankqualität in Kombination mit Erfahrungswerten der Projektteilnehmer.⁸⁵ Sie umfaßt fünf Hauptkategorien: telecommunications, retrieval software, data contents, search aids und costs.⁸⁶

Dem Qualitätsbegriff der Liste wurde dabei die Sichtweise eines Information professionals zugrunde gelegt. Qualitätsbestimmende Faktoren sind folglich, **validity** (Gültigkeit), **reliability** (Verlässlichkeit), **usuability** (Benutzbarkeit) in Verbindung mit **content** (Inhalt), **technical functioning** (technische Möglichkeiten) und generellem **ease of use** (Einfachheit in der Benutzung).⁸⁷

Die Kategorie **Telecommunications** beurteilt die Verbindung zum und die Kommunikation mit dem System. Dabei wird z.B. untersucht, ob die Verbindung häufigen Störungen unterliegt, ob die Anschaltzeiten der Datenbank im Handbuch oder Broschüren vermerkt sind oder wieviel Schritte nötig sind, um sich in die Datenbank einzuwählen.

Der Bereich **retrieval software** beschäftigt sich mit der Retrievalsprache und anderen technischen Aspekten des Systems sowie der Effektivität des

⁸² Vgl. Juntunen et. al. (1991), S. 352.

⁸³ Vgl. Juntunen et. al. (1995), S. 206.

⁸⁴ Vgl. Juntunen et. al. (1995), S. 207.

⁸⁵ Vgl. Juntunen et. al. (1995), S. 207.

⁸⁶ Vgl. Juntunen et. al. (1991), S. 353.

⁸⁷ Vgl. Juntunen et. al. (1995), S. 207.

Retrievalprogramms. In dieser Kategorie werden z.B. die Antwortzeiten der Datenbank, das Vorhandensein von logischen Operatoren, Thesauri oder SDI-Diensten beurteilt.⁸⁸ SDI (Selective Dissemination of Information) ist eine Bezeichnung für aktive Informationsdienste, bei denen einmalig ein Profil für Datenbankrecherchen zu einem speziellen Themengebiet festgelegt wird, die dann durch einen Host-Betreiber laufend durchgeführt und deren Ergebnisse dem Abonnenten des SDI-Dienstes regelmäßig geliefert werden.⁸⁹ Weiterhin werden in der Kategorie retrieval software die Ausgabemöglichkeiten einer Datenbank beurteilt, so z.B. die Ausgabeformate oder die Varianten der Dokumentenlieferung.⁹⁰

Die Kategorie **data contents** untersucht den Inhalt einer Datenbank. Kontrolliert werden sollen sowohl die Aktualität, die Genauigkeit und die Konsistenz der enthaltenen Daten als auch der Bereich von "coverage and scope". Im einzelnen werden z.B. die Häufigkeit der Updates, die Aktualität der Quellen und die Auswertungspraxis des Datenbankproduzenten geprüft. Das Vorhandensein von Abkürzungsverzeichnissen (Konsistenz) wird ebenfalls untersucht. Bezüglich der Genauigkeit kontrolliert das Komitee das Auftreten von "spelling errors".⁹¹

Der Bereich **search aids** untersucht die Hilfsmaterialien und Schulungsmöglichkeiten, die den Nutzern einer Datenbank zur Verfügung stehen. Die letzte Kategorie **costs** beurteilt die Kostenstruktur der Datenbank⁹².

Die Kriterien des Komitees zur Beurteilung der Qualität von Datenbanken berücksichtigen zwar die wichtigsten bei der Benutzung einer Datenbank auftretenden Qualitätsprobleme wie z.B. Schreibfehler, Ungenauigkeiten sowie Inkonsistenzen⁹³ und sind somit nach Meinung der Verfasserin grundsätzlich nachvollziehbar. Dennoch bedürfen verschiedene Ansatzpunkte der Kriterien der Diskussion. Die Anwendbarkeit der Kriterien wurde im Laufe ihrer Entwicklung zwar anhand acht verschiedener Datenbanken getestet, aber nicht jedes Kriterium konnte auf die jeweilige Datenbank angewandt werden. Dies dürfte darin begründet sein, daß die Liste der Qualitätskriterien keine Trennung nach Datenbanktypen vorsieht. Die Projektteilnehmer selbst stellten einen Schwerpunkt im Bereich bibliographischer Datenbanken fest.⁹⁴

⁸⁸ Vgl. Juntunen et. al. (1995), S. 216

⁸⁹ Die Begriffserklärung erfolgt in Anlehnung an Gaus (1995), S. 286.

⁹⁰ Vgl. Juntunen et. al. (1995), S. 216.

⁹¹ Vgl. Juntunen et. al. (1995), S. 217-218.

⁹² Vgl. Juntunen et. al. (1995), S. 219. Die vollständige Liste der Qualitätskriterien befindet sich im Anhang der Arbeit, vgl. unter VII.

⁹³ Vgl. Kapitel I, Abschnitt 1.1.

⁹⁴ Vgl. Juntunen et. al. (1991), S. 354.

Andere Datenbanktypen weisen jedoch Besonderheiten auf, denen durch eigene Kriterien Rechnung getragen werden sollte. In Volltextsystemen müßte dann z.B. auf die Suchbarkeit des gesamten Datensatzes geachtet werden

Weitere Kritikpunkte finden sich im Bereich der Kategorien **retrieval software** und **data contents** sowie im Ansatz des Komitees, den Kriterien die Sichtweise eines professionellen Informationsvermittlers zugrunde zu legen. In einer Zeit, in der zunehmend auch Endnutzer in Datenbanken recherchieren, die, meist nicht über ausreichende Fähigkeiten zur Beurteilung einer Datenbank verfügen,⁹⁵ ist es unbedingt notwendig, deren Anforderungen an eine Datenbank bei der Entwicklung von Beurteilungskriterien zu berücksichtigen.

In den Kategorien **retrieval software** und **data contents** werden mehrere qualitätssensible Bereiche wie Aktualität, "Coverage und Scope", Genauigkeit und Konsistenz gemeinsam beurteilt. In diesen Bereichen treten aber gerade die Fehler und Mängel auf, die bei einer Recherche mit besonders großer Wahrscheinlichkeit zum Verlust von relevanten Ergebnissen führen. Daher sollten sie ausführlicher und einzeln untersucht werden.

Außerdem sei noch erwähnt, daß die Liste des Komitees zur Beurteilung der Qualität von Datenbanken der Finnish Society for Information Services zwar auch außerhalb Finnlands bekannt ist, überwiegend jedoch nicht angewandt wird. Daraus könnte nach Meinung der Verfasserin auf eine mangelnde Benutzbarkeit der Kriterien insgesamt, geschlossen werden. Auch im Rahmen dieser Arbeit sollen die Kriterien des Komitees zur Beurteilung der Qualität von Datenbanken nicht angewandt werden.

1.3 Die Ansätze der Southern California Online User Group (SCOUG)

In den USA beschäftigte sich die Southern California Online User Group erstmals Ende der achtziger Jahre mit der Qualität von Datenbanken. Die SCOUG ist eine der ältesten Online User Groups. Sie wurde 1976 von ca. 50 Bibliothekaren, Online-Rechercheuren und anderen Personen aus dem Informationssektor gegründet.⁹⁶ Im Rahmen ihrer jährlichen Treffen nimmt die SCOUG zu aktuellen Themen aus den Arbeitsbereichen ihrer Mitglieder Stellung. 1990 lautete das Thema: „Measuring the

⁹⁵ Vgl. Kapitel I, Abschnitt 1.1.

⁹⁶ Vgl. Basch (1990 I), S. 252.

quality of data“. Ziel dieser Konferenz, deren Teilnehmer sich u.a. aus professionellen Online-Rechercheuren, Datenbankproduzenten, Vertretern von Hosts sowie Fachjournalisten zusammensetzten,⁹⁷ war die Entwicklung eines „Grundmodells“ zur Beurteilung der Qualität und Verlässlichkeit, einer „consumer rating scale“ für Datenbanken.⁹⁸ Die Konferenzteilnehmer beschäftigten sich mit drei verschiedenen Datenbanktypen: bibliographischen, Volltext- und Directory-Datenbanken. Das Grundmodell zur Qualitätsbeurteilung von Datenbanken sollte auf jeden Typ anwendbar sein. Es umfaßt die zehn folgenden Kategorien: „Consistency“, „Coverage/scope“, „Timeliness“, „Error rate/accuracy“, „Accessibility/Ease of use“, „Integration“, „Output“, „Dokumentation“, „Customer support and training“ und „Value-to-cost ratio“. Innerhalb des Modells wird keine Unterscheidung zwischen der eigentlichen Datenbank und dem sie bereitstellenden System getroffen.⁹⁹

Im Bereich der **Consistency** (Konsistenz) wird untersucht, ob die Feldanordnungen und –bezeichnungen sowie andere Datenelemente in einer konkreten Datenbank den gleichen Regeln folgen. In Volltextdatenbanken ist die Konsistenz ein Kriterium mit Bezug auf den Gesamteindruck der Datenbank.

In der Kategorie von **Coverage/scope** wird beurteilt, wie gut die Datenbank ihren jeweiligen Themenbereich bzw. ihr Fachgebiet abdeckt. Dazu muß die Indexierungspolitik des Datenbankproduzenten kontrolliert werden.

Der **Timeliness**-Begriff beschäftigt sich mit der Aktualität der Datenbank und untersucht die Häufigkeit der Updates oder die Aktualität des zugrundeliegenden Materials. Existieren verschiedene Versionen einer Datenbank, soll festgestellt werden, ob sich diese in ihrer Aktualität unterscheiden.

Im Bereich der **Accuracy/Error rate** wird die Genauigkeit und Fehlerrate einer Datenbank beurteilt. Dabei wird z.B. festgestellt, welche Quellen den Daten zugrunde liegen sowie deren Vollständigkeit geprüft. Weiterhin soll der prozentuale Fehleranteil in besonders wichtigen Feldern oder die Fehlerrate pro Update gemessen werden. Berücksichtigung finden aber auch die vom Produzenten der Datenbank installierten Maßnahmen zur Qualitätskontrolle.¹⁰⁰

Accessibility und **Ease of use** untersuchen sowohl den Zugang zum Datenbanksystem als auch zu den eigentlichen Daten. Dabei werden u.a. die Zugänglichkeit des Datenbanksystems beurteilt (gateway und/oder Direktzugang),

⁹⁷ Vgl. Tenopir (1990), S. 64.

⁹⁸ Vgl. Basch (1990 II), S. 18.

⁹⁹ Vgl. Basch (1990 II), S. 19.

aber auch die Möglichkeiten für den Abschluß eines Nutzungsvertrages (Möglichkeit der Online-Registrierung, wie z.B. beim SUBITO-Dienst des Deutschen Bibliotheksinstituts) untersucht. Zudem wird hier das Vorhandensein eines Online-Thesaurus sowie von Kontextoperatoren geprüft. Bezüglich der eigentlichen Daten soll die Tiefe und Flexibilität der Indexierung oder die Suchbarkeit verschiedener Elemente kontrolliert werden.

In der Kategorie **Integration** wird die Datenbank mit anderen Datenbanken desselben Typs aus dem gleichen Fachgebiet sowie mit Datenbanken des gleichen Datenbanksystems verglichen. Untersucht werden z.B. die Feldstruktur oder die Ausgabeformate. Außerdem fordert die SCOUG bei bibliographischen Datenbanken das Vorhandensein von Verknüpfungen zu Volltext-Datenbanken (im eigenen oder in fremden Datenbanksystemen). Beim Vergleich von Datenbanken desselben Systems wird darauf geachtet, ob eine Dublettenkontrolle installiert ist.

Im Bereich des **Output**-Begriffs sollen die Ausgabemöglichkeiten einer Datenbank beurteilt werden. Die SCOUG prüft z.B., ob Nutzer eigene Ausgabeformate definieren oder die Ergebnisse über ein Download in eigenen Systemen weiterverarbeiten können.¹⁰¹

Die Kategorie **Documentation** beschäftigt sich mit den zur Datenbank erhältlichen Hintergrund- und Hilfsmaterialien sowohl in gedruckter als auch in Online-Form. Geprüft wird z.B. das Vorhandensein regelmäßiger Newsletters, sonstiger Informationsmaterialien oder Online-Hilfen. Sind Online-Hilfen in der Datenbank installiert, wird getestet, ob diese kontextsensitiv sind. Untersucht wird auch, ob die in den Hintergrund- und Hilfsmaterialien enthaltenen Angaben bzgl. Indexierung, Aktualität usw. der Wahrheit entsprechen.

Der Bereich von **Customer support and Training** bewertet die den Nutzern zur Verfügung stehenden Schulungsmöglichkeiten sowie den „Kundendienst“.

In der Kategorie **Value-to-cost Ratio** unterliegt die Kostenstruktur der Datenbank einer generellen Beurteilung. Untersucht wird, ob das Datenbanksystem eine kosteneffiziente Suche zuläßt. Darüber hinaus werden die Preise nutzerdefinierter Ausgabeformate den voreingestellten Formaten gegenübergestellt.

¹⁰⁰ Vgl. Basch (1990 II), S. 19.

¹⁰¹ Vgl. Basch (1990 II), S. 20.

In Ergänzung dieses Grundgerüsts hat die SCOUG drei weitere Modelle zur Beurteilung von bibliographischen, Volltext- und Directory-Datenbanken entwickelt.¹⁰²

Bibliographische Datenbanken sollten u.a. folgende Eigenschaften aufweisen:

- ?? Möglichkeit, nach Land, Sprache oder Fachbegriffen (subject facet) einzuschränken;
- ?? exakte Wiedergabe von Autor und Titel;
- ?? Möglichkeit der Suchbarkeit der „author-affiliation“, inklusive vollständiger Adresse;
- ?? Verwendung eines Standard-Formats zur Erfassung der Autorennamen;
- ?? verzichten auf Abkürzungen in den „Source-Feldern“ (Quelle), insbesondere bei Zeitschriften sowie
- ?? Möglichkeit der grundsätzlichen Suchbarkeit aller Felder.¹⁰³

Für **Volltextdatenbanken** gelten ebenfalls besondere Anforderungen:

- ?? Die Suchbarkeit des gesamten Datensatzes im Freitext, aber auch die Suchbarkeit nach Feldern sollte möglich sein;
- ?? eine automatische Suche nach Synonymen oder der Pluralform eines Suchbegriffs sollte durchführbar sein;
- ?? Volltextsysteme sollten die Möglichkeit des „Keyword in Context“ (KWIC) nutzen und
- ?? bzgl. „Coverage and Scope“ ihre Indexierungspolitik bekanntgeben. Lücken in der Auswertung oder die Aufgabe der Auswertung bei einzelnen Quellen müssen mitgeteilt werden.¹⁰⁴

Im Bereich der **Directory-Datenbanken** sind der Zugang zum Datenbanksystem und der Inhalt, also die eigentlichen Daten, besonders wichtig. Die SCOUG fordert für diese Systeme:

- ?? generelles Vorhandensein von Basisinformationen in Form von Ansprechpartner und Adressen unabhängig vom Typ der Directory-Datenbank;

¹⁰² Vgl. Basch (1990 II), S. 21.

¹⁰³ Vgl. Basch (1990 II), S. 21-22.

¹⁰⁴ Vgl. Basch (1990 II), S. 22.

- ?? Angabe der Herkunft und des Stands der Informationen; Löschung nicht mehr existierender Firmen;
- ?? Angaben zur Währung bei Finanzdaten, bei Nicht-US-Währungen auch Angabe des Tags der letzten Aufwertung;
- ?? Herstellung von Verbindungen zwischen Mutter- und Tochtergesellschaften.¹⁰⁵

In Ergänzung zu den generellen Kriterien des Grundmodells sowie den speziellen Kriterien zur Beurteilung der drei genannten Datenbanktypen entwickelte die SCOUG noch eine Liste mit Anforderungen an Endnutzer-Systeme. Die Anforderungen eines Endnutzers einer Datenbank unterscheiden sich von denen eines professionellen Informationsvermittlers. Endnutzer benötigen in der Regel keine umfassende und vollständige Recherche; ein Ergebnis mit wenigen guten Artikeln zum gesuchten Themenkomplex ist ausreichend.¹⁰⁶

Endnutzerorientierte Systeme sollten u. a.:

- ?? Hilfe beim Auswählen einer geeigneten Datenbank und beim Formulieren der Suchanfrage geben;
- ?? kontextsensitive Hilfefunktionen besitzen;
- ?? verschiedene alternative Lieferformen für die Dokumente anbieten;
- ?? die Preise für die Recherche nach der Anzahl der Ergebnisse bestimmen.¹⁰⁷

Die durch die SCOUG entwickelten Kriterien sollten ursprünglich z.B. mit Hilfe einer Zehn-Punkte-Skala gewichtet werden. Dabei sollte zunächst jedes Kriterium mit Punkten bewertet und die Teilergebnisse anschließend zu einem Gesamtergebnis der entsprechenden Datenbank zusammengefaßt werden.¹⁰⁸ Diese Gewichtung und eine Bewertung mit Punkten scheint jedoch nicht erfolgt zu sein.

Der Ansätze der SCOUG stießen in der Fachwelt auf große Resonanz und Akzeptanz. Bereits 1991 bildeten ihre Qualitätskriterien die Grundlage einer Umfrage der NFAIS (National Federation of Abstracting and Information Services)

¹⁰⁵ Vgl. Basch (1990 II), S. 22.

¹⁰⁶ Vgl. Basch (1990 II), S. 22.

¹⁰⁷ Vgl. Basch (1990 II), S. 22-23.

¹⁰⁸ Vgl. Basch (1990 II), S. 23. Die vollständige Liste der Kriterien findet sich im Anhang der Arbeit unter VII.

bei Datenbankproduzenten bezüglich Qualitätssicherungsmaßnahmen. Auch 1993/1994 im Rahmen des EQUIP-Projektes wurden sie verwendet.¹⁰⁹ Einzelne Datenbankproduzenten, z.B. *Granick*, setzten sich gleichfalls mit den Modellen der SCOUG auseinander.¹¹⁰

Drei wichtige Datenbanktypen (bibliographische, Volltext- und Directory-Datenbanken) werden durch spezielle Kriterien beurteilt, wodurch die Besonderheiten der jeweiligen Datenbanktypen nach Meinung der Verfasserin besser als im finnischen Modell berücksichtigt werden. Durch die getrennte Untersuchung der unterschiedlichen Qualitätsmängel in den Kategorien **consistency**, **accuracy/error rate**, **Coverage/Scope** und **Timeliness** wird deren Wichtigkeit Rechnung getragen sowie eine größere Übersichtlichkeit und Ausführlichkeit in der Beurteilung erreicht. Hervorgehoben werden muß auch die Berücksichtigung der Endnutzerbedürfnisse. Am gesamten Ansatz der SCOUG kann somit die intensive Auseinandersetzung mit den Fehlern und Qualitätsmängeln einer Datenbank abgelesen werden. Daher wird er auch bereits erfolgreich angewendet. Das Centre for Information Quality Management (CIQM) machte sie zur Grundlage seiner Arbeit, die im folgenden dargestellt werden soll.

1.4 Die Ansätze des Centre for Information Quality Management (CIQM)

In Großbritannien beschäftigt sich das Centre for Information Quality Management (CIQM), seit seiner Gründung durch die Library Association und die UK Online User Group im Jahre 1993, mit der Qualität von Datenbanken. Anstelle der beiden Gründungsmitglieder werden die Tätigkeiten des CIQM heute durch Information Automation Limited ausgeführt.

Das CIQM versteht sich als Clearinghouse, an das bei der Nutzung von Datenbanken (Online oder CD-ROM) aufgetretene Fehler oder Qualitätsprobleme gemeldet werden können. Diese Qualitätsmängel werden anschließend durch das CIQM an jene Stelle im Bereich des Informationssektors weitergegeben, die für den Fehler verantwortlich ist. Dies können der Provider, der Host oder auch die Produzenten einer CD-ROM

¹⁰⁹ Vgl. IMO working paper 95/4 (<http://www2.echo.lu/impact/imo/9504.html>); Wilson (1998), S. 345-357.

¹¹⁰ Vgl. Granick (1991), S. 117-136.

sein. Die Antwort der zuständigen Stelle wird dann

wiederum an den Nutzer übermittelt. Alle mit diesem Prozeß verbundenen Arbeiten sind für die Nutzer einer Datenbank grundsätzlich kostenfrei.

Das CIQM hat die durch die Nutzer von Datenbanken gemeldeten Mängel genutzt, um einen Ansatz zu entwickeln, durch den die Qualität einer Datenbank festgestellt werden kann, die sogenannten Database Labels, denen die durch die SCOUG entwickelten Modelle zur Qualitätsbeurteilung zugrunde liegen.¹¹¹

Database Labels wurden erstmals 1993 durch Jascó vorgeschlagen.¹¹²

Sie führen, so *Armstrong*, zu einer Abkehr vom bis zum Zeitpunkt der Entwicklung der Labels üblichen, WYGIWIG (What You Get Is What Is Given), zum WYSIWYG (What You See Is What You Get).¹¹³

Jedes Database Label stellt eine Minispezifikation mit einem Standardaufbau dar, die daher auch Vergleiche zwischen verschiedenen Datenbanken zuläßt. Durch die Definition der Labels als Minispezifikation ist gleichzeitig festgelegt, daß jede Version einer Datenbank ein eigenes Label erhalten muß.¹¹⁴

Ein Database Label besteht aus vier Ebenen. Die erste Ebene (**General Description**) enthält Aussagen des Datenbankproviders zu seiner Auswertungs- und Indexierungspolitik, Angaben zu "coverage and scope", dem Beginn des Berichtszeitraums sowie zur Durchführung von Updates. Die zweite Ebene (**Quality Assurance Policy Statements**) setzt sich aus Aussagen zur Qualitätssicherung der Datenbank zusammen. Diese entstehen z.T. in Zusammenarbeit mit dem Verkäufer/Anbieter der Datenbank. Die dritte Kategorie beschreibt mit Hilfe von numerischen Angaben die Struktur der Datenbank (**Database Structure**). Bestandteil dieser Beschreibung sind Aussagen zu suchbaren und nicht-suchbaren Feldern sowie garantierten Feldformaten. Die vierte Ebene (**Coverage and General Information**) enthält prozentuale Angaben zur "geographic coverage" und ergänzende Aussagen z.B. zur Software und den Ansprechpartnern beim Provider.¹¹⁵

Labels müssen so veröffentlicht werden, daß sie für die Nutzer einer Datenbank einfach erreichbar sind. *Armstrong* schlägt hierfür das Internet vor.¹¹⁶ Eine weitere

¹¹¹ Vgl. *Armstrong* (1994), S. 28.

¹¹² Vgl. Jascó (1993 II), S. 7-9.

¹¹³ Vgl. *Armstrong* (1994), S. 28-29.

¹¹⁴ Vgl. *Armstrong* (1999) <http://www.i-a-l.co.uk/ciqm/BULISC.htm>.

¹¹⁵ Der Aufbau der Labels ist einem über die Homepage des CIQM unter <http://www.la-hq.org.uk/liaison/ciqm/labels.htm> einsehbaren Beispiel-Label entnommen.

¹¹⁶ Vgl. *Armstrong* (1999) <http://www.i-a-l.co.uk/ciqm/BULISC.htm>.

Möglichkeit wäre, den üblichen Materialien, z.B. der gedruckten Dokumentation einer Datenbank, die Labels beizulegen.¹¹⁷

Durch die Database Labels sind Nutzer einer Datenbank in der Lage, die Qualität eines Produktes festzustellen, bevor sie es erwerben oder zur Recherche nutzen.¹¹⁸

Database Labels bieten Informationen über die in einer konkreten Datenbank enthaltenen Informationen, stellen deshalb Metainformationen dar. Was aus den bereits beschriebenen Gründen vor allem für Endnutzer besonders wichtig ist.

Da das Centre for Information Quality Management diesen Labels keine eigenen, sondern die Kriterien der SCOUG zugrunde legt soll der Ansatz im Rahmen dieser Arbeit nicht weiterverfolgt werden. Dennoch ist die Verfasserin der Ansicht, daß mit den Labels ein Ansatz entwickelt wurde, der in knapper und gut strukturierter Form die wichtigsten qualitätssensiblen Bereiche einer Datenbank darstellt und zudem allgemein verständlich ist. Die weitere Nutzung der Labels sollte daher unbedingt diskutiert werden.

Database Labels könnten dabei auch als eine Art Zertifizierung von Datenbanken verstanden werden, die auch in Deutschland erfolgen könnte. Hierbei wäre zu beachten, daß eine aussagekräftige Zertifizierung von Datenbanken, nur durch fachlich qualifizierte und von den Datenbankproduzenten und -anbietern unabhängige Organisationen erfolgen kann. Dies könnten nach Meinung der Verfasserin z.B. Einrichtungen des Bibliothekswesens, ähnlich dem CIQM, sein. Derartige Einrichtungen bringen eine fachliche Qualifizierung quasi automatisch mit sich, da ihre tägliche Arbeit durch die permanente Auseinandersetzung mit Datenbanken gekennzeichnet ist. Außerdem sind diese Einrichtungen durch ihre meist staatliche Finanzierung von Datenbankproduzenten und -anbietern unabhängig, so daß die für eine Zertifizierung notwendige Distanz vorhanden wäre.

¹¹⁷ Vgl. Armstrong (1994), S. 28. Ein beispielhaftes Database Label findet sich im Anhang der Arbeit unter VII.

¹¹⁸ Vgl. Armstrong (1994), S. 28.

2. Ansätze zur Qualitätsbeurteilung im Bereich des Internet

Im Bereich des Internet wurden bereits zahlreiche Ansätze zur Qualitätsbeurteilung entwickelt. Diese beziehen sich jedoch meist nicht ausdrücklich auf die Qualität von Datenbanken sondern auf die Qualität einzelner Internet-Dokumente. Im folgenden ist daher zu prüfen, welcher der vorhandenen Ansätze auf die Beurteilung von Internet-Datenbanken übertragen, insbesondere zur Qualitätsbeurteilung der beiden CORDIS-Datenbanken genutzt werden kann.

2.1 Die WWW Virtual Library

Die WWW Virtual Library war eines der ersten Projekte, das sich mit der Erschließung von Internet-Dokumenten beschäftigte.¹¹⁹ Weltweit werden von Freiwilligen im Netz Informationen zu verschiedenen Themengebieten gesammelt und zur Verfügung gestellt. Grundsätzlich wurden keine einheitlichen Kriterien bezüglich der Dokumentenauswahl, sondern lediglich ein Grundmodell bezüglich des Aufbaus einer Virtual Library entwickelt.¹²⁰

Die Kriterien des Grundmodells wurden in einer eigenen Information Quality WWW Virtual Library erfaßt. Als Anforderungen werden hierfür festgelegt: "*provide their own information*", "*are usefull and informative*", "*are easy to find*", "*are universally accessible*", "*are well structured and organised*", "*are well formatted and edited*" sowie "*are easy to establish, run, maintain and improve on*".¹²¹ Die Auswahlkriterien einzelner WWW Virtual Libraries sind nicht veröffentlicht.

Obwohl die Vorschriften zum Aufbau einer WWW Virtual Library auch Ansätze bezüglich einer Bewertung ihres Inhalts enthalten („*are usefull and informative*“), scheint die Gestaltung der Verzeichnisse im Vordergrund zu stehen. Die Verfasserin ist demgegenüber der Ansicht, daß gerade der Inhalt einer Internet-Datenbank, z.B. bezüglich der Genauigkeit, Verlässlichkeit oder Aktualität der verzeichneten Daten, ihre Qualität entscheidend beeinflusst. Die Gestaltung stellt einen zweitrangigen Aspekt dar, denn ohne den entsprechenden Inhalt ist die beste Gestaltung wertlos. Der

¹¹⁹ Vgl. Payer (1997) <http://machno.hbi-stuttgart.de/?payer/infoq.html>.

¹²⁰ Vgl. Payer (1997) <http://machno.hbi-stuttgart.de/?payer/infoq.html>.

¹²¹ Vgl. Ciolek (1997) <http://www.ciolek.com/WWWVLPages/QtyPages/QtyTruisms.html>.

Ansatz der WWW Virtual Library soll daher im Rahmen dieser Arbeit nicht weiterverfolgt werden.

2.2 Die Ansätze der "Internet review sites"

Weitere Ansätze zur Qualitätsbeurteilung von Internet-Dokumenten finden sich in den zahlreichen Seiten, welche sich der Beurteilung von Internet-Ressourcen verschrieben haben. Die von diesen sogenannten "Internet review sites" verwendeten Kriterien können stark variieren.¹²² Da nur wenige Internet review sites ihre Bewertungskriterien offenlegen¹²³ und die veröffentlichten Ansätze zudem sehr fragwürdig sind, soll im folgenden lediglich auf zwei Beispiele eingegangen werden.

Yahoo! Internet Life benutzt z.B. ein fünfstufiges Rating-System, bei dem die folgenden Bewertungen vergeben werden: "Pointless, don't bother", "nothing special, same old", "good, functional, not unique", "top-notch, well above average" und schließlich "indispensable, a must-see".¹²⁴ Diese Kriterien sind nach Ansicht der Verfasserin zweifelhaft. Mit Feststellungen wie "*good functional, not unique*" oder "*indispensable, a must-see*" lassen sich keine Aussagen über die wirkliche Qualität eines Internet-Dokuments oder einer -Datenbank machen.

WebCrawler listet nach eigener Aussage nur die besten Seiten auf. Es wird weiter festgestellt: "*These sites are chosen for their depth of content, excellent organization and timeliness.*"¹²⁵ Eine nähere Definition der Begriffe "beste Seite" oder "excellent organization" wird jedoch nicht gegeben.

Im Bibliotheks- und Informationswesen sind die Ansätze der Internet review sites daher zurecht starker Kritik ausgesetzt. Cooke stellt z.B. fest: "*However, such services are themselves of varying usefulness and quality.*"¹²⁶ Rettig¹²⁷ und Collins¹²⁸ lehnen die Ansätze zur Qualitätsbeurteilung der Internet reviews sites gleichfalls ab.

Eine Anwendung dieser Kriterien zur Beurteilung der Qualität der beiden CORDIS-Datenbanken soll daher im Rahmen dieser Arbeit nicht erfolgen.

¹²² Vgl. Armstrong (1997) http://www.i-a-l.ci.uk/ciqm/db_qual/2_1_4.htm; ähnlich Rettig (<http://www.onlineinc.com/onlinemag/septOL/rettig9.html>) oder Cooke et. al. (1996), S. 35.

¹²³ Vgl. Cooke et. al. (1996), S. 35.

¹²⁴ Die Auswahlkriterien von Yahoo! Internet Life finden sich unter <http://www.zdnet.com/yil/content/depts/stargde.html>.

¹²⁵ <http://webcrawler.com/Help/Guide.html>.

¹²⁶ Cooke et. al. (1996), S. 34.

¹²⁷ Vgl. Rettig (1996) unter <http://www.onlineinc.com/onlinemag/SeptOL/rettig9.html>.

¹²⁸ Vgl. Collins (1996), S. 124.

2.4 Ansätze zur Qualitätsbeurteilung des Internet innerhalb des Bibliotheks- und Informationswesens

Im Bibliotheks- und Informationswesens wurden zahlreiche eigene Ansätze zur Qualitätsbeurteilung von Internet-Dokumenten entwickelt. Diese sind dadurch gekennzeichnet, daß sie z.B. versuchen, die Kriterien, welche in Bibliotheken bei der Auswahl von Printmaterialien benutzt werden, auf den Bereich des Internet zu übertragen. Andere Modelle orientieren sich an den Qualitätskriterien, die zur Beurteilung von Online-Datenbanken verwendet werden.¹²⁹ Im folgenden soll beispielhaft auf wenige Ansätze eingegangen werden.

Ein Vertreter des ersten Modells ist Collins, der im einzelnen die Kategorien „Content“, „Authority“, „Currency“, „Organization“, „Accessibility“ und „Search engine“ zur Qualitätsbeurteilung vorschlägt.¹³⁰

Mit der **Content**-Kategorie soll der Inhalt eines Dokuments beurteilt werden. Im Bereich der **Authority** wird geprüft, ob ein Autor im entsprechenden Fachgebiet bereits einen Namen hat und die Quelle damit vertrauenswürdig ist. Die Kategorie der **Currency** soll zur Aktualitätsbeurteilung genutzt werden, während die Bereiche **Organization** und **Accessibility** den Aufbau und die Zugänglichkeit einer Internet-Ressource bewerten. Durch den **search engine** soll innerhalb des jeweiligen Dokuments eine Recherche möglich sein.¹³¹

Den gleichen Ansatz vertritt auch *Rettig*, der allerdings eine etwas andere Aufteilung vornimmt und von insgesamt achtzehn verschiedenen Kriterien ausgeht. Er legt **Accuracy** (Genauigkeit), **Appropriateness** (Angemessenheit in Bezug auf die Zielgruppe), **Arrangement** (Gestaltung), **Authority**, **Bibliography** sowie **Comparability** (Vergleichbarkeit mit anderen Dokumenten) als Kategorien zur Qualitätsbeurteilung fest. Darüber hinaus führt er **Completeness** (Vollständigkeit), **Content**, **Distinction** (Abweichungsmerkmale einer Ressource von einer anderen), **Documentation**, **Durability** (Dauerhaftigkeit), **Ease of use**, **Illustration**, **Index** (Vorhandensein des von Collins beschriebenen search engines) sowie **Level**, **Reliability** (Verlässlichkeit), **Revisions** (Häufigkeit der Updates) und **Uniqueness** (Einmaligkeit) als zusätzliche Kriterien an.¹³²

¹²⁹ Siehe dazu 1.3 in diesem Kapitel.

¹³⁰ Vgl. Collins (1996), S. 124.

¹³¹ Vgl. Collins (1996), S. 124.

¹³² Vgl. Rettig (1996) unter <http://www.onlineinc.com/onlinemag/SeptOL/rettig9.html>.

An den Ansätzen zur Qualitätsbeurteilung von Online-Datenbanken orientiert sich z.B. *Armstrong*, der die Vergabe der CIQM-Labels auch für Internet-Dokumente vorschlägt.¹³³

Auch die SCOUG hat sich im Rahmen ihrer neunten Jahrestagung mit der Qualität von Internet-Ressourcen beschäftigt und insgesamt 20 Kriterien hierfür formuliert. Ausgangspunkt waren dabei gleichfalls ihre Kriterien zur Beurteilung von Online-Datenbanken. Gefordert werden u.a. „**Credibility**“ (Glaubwürdigkeit), „**Authority**“ (Autorität), „**Indexing**“ (konsistente Indexierung), „**Registration**“ (möglichst automatische Identifikation des Urheber eines Dokuments), „**Reviews / Ratings**“ (Übersichten und Bewertungen analog den Klappentexten in Büchern) aber auch „**Security and Privacy**“ (Datenschutz), „**Feedback / Maintenance / Customer service**“ (Möglichkeiten zur Kommunikation mit dem Nutzer), „**Copyright and protection of intellectual property**“ (Beachtung des Copyright und der Schutz geistigen Eigentums) sowie „**maintenance**“ (angemessene Instandhaltung).¹³⁴

Die soeben beschriebenen Ansätze enthalten sowohl Kriterien bezüglich des Inhalts, wie z.B. "Accuracy", "Appropriateness" oder "Completeness" als auch bezüglich der Struktur ("Arrangement" oder "Ease of use") einer Internet-Ressource. Sie untersuchen somit auch jene Bereiche, die nach der hier vertretenen Ansicht qualitätsentscheidend sind, und wären daher grundsätzlich zur Qualitätsbeurteilung der beiden CORDIS-Datenbanken geeignet. Bezüglich der Benutzbarkeit ist jedoch festzustellen, daß mit Ausnahme des Ansatzes von *Collins* keiner dieser Ansätze erfolgreich angewandt wird.¹³⁵ Auch im Rahmen der hier vorliegenden Arbeit sollen diese Ansätze daher nicht angewandt werden.

¹³³ Vgl. *Armstrong* (1997) http://www.i-a-l.co.uk/ciqm/db_qual/2_1_4.htm.

¹³⁴ Vgl. *Ebbinghouse* (1995), S. 48-56.

¹³⁵ *Collins* verwendet die beschriebenen Kriterien im Rahmen des sogenannten Infofilter-Projekts, einer Bewertung von Internet-Seiten im Internet, zu finden unter <http://www.usc.edu/user/help/flick/Infofilter>.

2.4 The Argus Clearinghouse

The Argus Clearinghouse wurde 1993 an der University of Michigan School of Information and Library Studies gegründet. Ziel von Argus ist die Zusammenstellung von qualitativ hochwertigen "guides", für die man zentrale Sammelstelle sein möchte. Ein "guide" stellt dabei ein über das Internet zugängliches, kostenfreies auf ein bestimmtes Thema bezogenes Verzeichnis dar, in dem weitere Quellen aus dem Internet verzeichnet und beschrieben werden.

The Argus Clearinghouse recherchiert nicht selbst, sondern bewertet lediglich die "guides", die durch Außenstehende (meist durch die Urheber der guides) gemeldet werden. Die Beurteilung erfolgt mit Hilfe eines selbst entwickelten Rating-Systems, das fünf Ebenen aufweist: "Level of Ressource Description", "Level of Ressource Evaluation", "Guide Design", "Guide Organizational Schemes" sowie "Guide Meta Information". Die gemeldeten guides werden nur beurteilt, wenn sie folgende Voraussetzungen erfüllen: Angabe des Autorennamens, Arbeitsstelle des Autors, Angabe einer Email-Adresse, Kurzbiographie des Autors, Angabe der letzten Update-Erstellung des guides.

Die Ebene des **Level of Ressource Description** untersucht die Beschreibung eines Dokuments im jeweiligen Guide, der **Level of Ressource Evaluation** prüft die durch die guides vorgenommene Bewertung ihrer Seiten. Dabei wird z.B. beurteilt, inwieweit ein guide die Aktualität oder den Inhalt seiner Quellen untersucht hat. Die Bereiche **Guide Design** und **Guide Organizational Scheme** beurteilen das Design und die Anordnung des Verzeichnisses. Die Kategorie **Guide Meta Information** untersucht die vorhandenen Informationen über den guide (Metainformationen). Geprüft wird z.B., ob Informationen zum Autor und seiner fachlichen Qualifikation oder auch zur Indexierungspolitik des guides vorhanden sind.¹³⁶ Die Bewertungen aus allen fünf Ebenen werden zu einem Gesamtergebnis zusammengefaßt, das in Form von maximal fünf Haken dargestellt wird.

The Argus Clearinghouse untersucht und bewertet fachliche Verzeichnisse von Internetdokumenten. Da diese Verzeichnisse auf bestimmte Ressourcen im Netz verweisen, sind sie nach Meinung der Verfasserin dieser Arbeit einer Datenbank grundsätzlich vergleichbar. Eine Anwendung des Ansatzes von The Argus Clearinghouse zur Qualitätsbeurteilung der CORDIS-Datenbanken wäre also

¹³⁶ Sämtliche Informationen wurden der Homepage von "The Argus Clearinghouse" unter <http://www.clearinghouse.net> entnommen.

grundsätzlich möglich. Hinzukommt, daß die durch The Argus Clearinghouse verwendeten Kriterien jene Bereiche eines Verzeichnisses beurteilen, welche für die Qualität ausschlaggebend sind. Dies würde für ihre Anwendung zur Qualitätsbeurteilung der beiden CORDIS-Datenbanken sprechen. Die Verfasserin ist allerdings der Ansicht, daß die Darstellung der Kriterien in Form von Haken zu wenig aussagekräftig ist. Um eine Vorstellung vom Inhalt des guides zu erhalten, ist es trotz erfolgter Bewertung notwendig, sich den eigentlichen guide anzusehen. Ausformulierte Bewertungen wären wesentlich hilfreicher, da sie zusätzliche Informationen enthielten, die Rückschlüsse auf die Bewertung des guides zulassen würden.

2.5 Die Ansätze zur Qualitätsbeurteilung des DESIRE-Projekts

Die aktuellste Entwicklung zur Qualitätsbeurteilung von Internet-Dokumenten ist das DESIRE-Projekt. DESIRE (Development of a European Service for Information and Research in Education) ist ein EU-finanziertes, internationales Projekt mit dem Ziel, Wissenschaftlern im Internet ein Informationsnetzwerk zur Verfügung zu stellen. Dazu beschäftigen sich die DESIRE-Projektteilnehmer sowohl mit dem Auffinden relevanter Informationen, als auch mit der Datensicherheit, der langfristigen Speicherung dieser Informationen sowie mit Maßnahmen zur Nutzerschulung.¹³⁷

Im Rahmen eines Teilprojektes wird versucht, die Katalogisierung und Indexierung von Internet-Dokumenten durch die Weiterentwicklung von sogenannten "subject gateways" zu verbessern. Ein **subject gateway** ist ein themenbezogenes Verzeichnis von Internet-Ressourcen, das den Einstieg in ein bestimmtes Fachgebiet erleichtern soll. DESIRE greift auf die im Rahmen nationaler Projekte entstandenen gateways seiner Partnerorganisationen und die gateways des UK Electronic Libraries Programme zurück.¹³⁸ Im einzelnen sind dies:

?? ADAM (Art, Design, Architecture & Media gateway),

?? BUBL (The Bulletin Board for Libraries),

?? EELS (Engineering Electronic Library),

¹³⁷ Vgl. Hiom (1998) <http://www.shef.ac.uk/?is/publications/infres/paper48.html>

¹³⁸ Vgl. Hofman / Worsfold (1998) http://www.ukoln.ac.uk/metadata/desire/quality/report_1.html.

- ?? EEVL (Edinburgh Engineering Virtual Library),
- ?? NBW (Nederlandse Basisclassificatie Web),
- ?? OMNI (Organising Medical Networked Information),
- ?? ROADS (Ressource Organisation And Discovery in Subject-based services),
- ?? RUDI (Ressource for Urban Design Information) und
- ?? SOSIG (Social Science Information Gateway).¹³⁹

Als Teilziel des Katalogisierungs- und Indexierungsprojektes wurden Kriterien zur Auswahl qualitativ hochwertiger Internet-Ressourcen entwickelt. Dazu trugen die Projektteilnehmer sowohl die von den subject gateways verwendeten Auswahlmodelle als auch die in der Literatur diskutierten Ansätze zur Qualitätsbeurteilung von Online-Datenbanken und Internet-Dokumenten zusammen. Zusätzlich fanden die in der Industrie genutzten Kriterien zur Qualitätsmessung Berücksichtigung. Diese außerhalb des DESIRE-Projektes verwendeten Ansätze wurden durch die Projektteilnehmer zusammengefaßt, bewertet und bzgl. ihrer Anwendbarkeit immer wieder getestet.¹⁴⁰

Schließlich entstanden die Selection criteria for quality controlled information gateways, welche aus folgenden fünf Hauptkategorien bestehen: der „Scope Policy“, den „Content Criteria“, den „Form Criteria“, den „Process Criteria“ sowie der „Collection Management Policy“.¹⁴¹

Im Bereich der **Scope Policy** werden Grundüberlegungen bzgl. der Auswahl der Quellen angestellt, wobei die Informationsbedürfnisse potentieller Nutzer berücksichtigt werden. Festgelegt werden u.a. das Thema, der geeignete Dokumenttyp sowie der Schwierigkeitsgrad einer Ressource. Zusätzlich werden die Zugänglichkeit zu den Quellen, mögliche Kosten, die Technologie oder Sicherheitsaspekte bedacht. Im Bereich der Scope Policy sollen auch die Katalogisierungspolitik, die Sprache sowie der geographische Raum, aus dem die Quellen stammen, fixiert werden.

Mit den **Content Criteria** soll der Inhalt einer Internet-Ressource beurteilt werden. In dieser Kategorie wurden die in Bibliotheken benutzten Kriterien zur Auswahl von Printmaterialien verarbeitet. Untersucht wird die Gültigkeit der Information (Validity), die Autorität und Reputation des Dokuments bzw. des Autors

¹³⁹ Vgl. Hofman / Worsfold (1998) <http://www.ukoln.ac.uk/metadata/desire/quality/toc.html>.

¹⁴⁰ Vgl. Hofman / Worsfold (1998) <http://www.ukoln.ac.uk/metadata/desire/quality/report-1.html>.

¹⁴¹ Zu den Auswahlkriterien vgl.: Hofman / Worsfold (1998) unter http://www.ukoln.ac.uk/metadata/desire/quality/report_2.html.

(Authority and Reputation), die Stichhaltigkeit des Dokuments (Substantiveness), seine Genauigkeit (Accuracy), Vollständigkeit/Ausführlichkeit (Comprehensiveness) sowie seine Einmaligkeit (Uniqueness) und sein Aufbau (Composition).

Die Ebene der **Form Criteria** setzt sich mit der Organisation und Gestaltung einer Quelle auseinander. Geprüft wird u.a. die Einfachheit der Navigation und das Angebot von Nutzerunterstützung in Form von Hilfefunktionen oder Dokumentationsmaterialien. Außerdem wird die Verwendung anerkannter Standards sowie einer angemessenen Technologie kontrolliert.

Der Bereich der **Process Criteria** beurteilt die Aspekte "Information Integrity", "Site Integrity" und "System Integrity". Geprüft wird die Aktualität und Dauerhaftigkeit der Quelle wie auch der darin enthaltenen Informationen. Darüber hinaus soll die technische Leistungsfähigkeit und die Stabilität des Systems untersucht werden.

Die **Collection Management Policy** bewertet eine Internet-Ressource im Umfeld des konkreten subject gateways. Untersucht werden die Bereiche "Collection Coverage and Balance", "Availability of Internet-Resources" und "Availability of Library Resources". Geprüft wird der Wert einer Quelle in bezug zu anderen bereits im gateway verzeichneten Dokumenten. Dabei werden auch das Vorhandensein anderer Internet-Dokumente zum gleichen Thema oder die Existenz von anderen Informationsmaterialien (z.B. in Print) berücksichtigt.¹⁴²

Die im Rahmen des Projektes entwickelten Auswahlkriterien¹⁴³ nehmen zu allen qualitätsentscheidenden Bereichen eines Internet-Dokuments Stellung. Sie stellen nach Meinung der Verfasserin eine Zusammenfassung und Weiterentwicklung aller bisherigen Ansätze zur Qualitätsbeurteilung im Bereich des Internet dar. Eine einheitliche Struktur in den Kriterien ist erkennbar. Auch nach erfolgreichem Projektabschluß werden die Selection Criteria for Quality Controlled Information gateways weiterhin zur Auswahl von qualitativ hochwertigen Internet-Ressourcen verwendet. Von einer Benutzbarkeit des Ansatzes kann daher ausgegangen werden. Obwohl sie eigentlich zur Auswahl qualitativ hochwertiger Internet-Ressourcen für subject gateways entwickelt wurden, können die Selection criteria for quality controlled information gateways nach Meinung der Verfasserin auch zur Qualitätsbeurteilung von Internet-Datenbanken verwendet werden. Ein subject gateway stellt eine Sammlung von Daten/Dokumenten dar und ist insofern einer Datenbank durchaus vergleichbar.

¹⁴² Vgl. Hofman / Worsfold (1998) unter http://www.ukoln.ac.uk/metadata/desire/quality/report_2.html.

IV. Darstellung und Begründung der gewählten Kriterien zur Qualitätsbeurteilung

1. Kriterien zur Beurteilung der SIGLE-Datenbank

Zur Beurteilung der SIGLE-Datenbank soll sechs der insgesamt zehn Kriterien des SCOUG-Modells verwendet werden.

In zwei durch EUSIDIC und das CIQM durchgeführten unabhängigen Umfragen unter Informationsfachleuten wurden

- ?? Consistency (Konsistenz),
- ?? Coverage and Scope (Themenbereich/Fachgebiet),
- ?? Timeliness (Aktualität der Datenbank und der verzeichneten Quellen),
- ?? Accessibility/Ease of use (Zugang und Benutzbarkeit),
- ?? Accuracy/Error rate (Genauigkeit und Fehlerrate)

als wichtigste Kriterien zur Qualitätsbeurteilung einer Datenbank genannt¹⁴⁴. Diese Kriterien sollen daher auch im Rahmen dieser Arbeit zur Anwendung kommen. Hinzukommen soll die Kategorie „Documentation“, da sie nach Meinung der Verfasserin in unmittelbarem Bezug zu den Kriterien „Coverage and Scope“ sowie „Consistency“ steht. Beide Kategorien können ohne eine Heranziehung des entsprechenden Dokumentationsmaterials der SIGLE-Datenbank nicht bewertet werden.

Durch die gewählten Kriterien ist eine Untersuchung der Bereiche der SIGLE-Datenbank möglich, in denen Fehler am wahrscheinlichsten sind.¹⁴⁵ Die übrigen Kriterien des SCOUG-Modells, z.B. „Value-to-Cost Ratio“ oder „Customer support and Training“, sind demgegenüber nach Meinung der Verfasserin zweitrangig. Ein angemessenes Kosten-Nutzen-Verhältnis (Value-to-cost Ratio) z.B. kann nur dann vorliegen, wenn die Datenbank eine geringe Fehlerrate aufweist, aktuell und konsistent ist.

Zusätzlich soll die Endnutzerorientierung der SIGLE-Datenbank getestet und das Modell zur Beurteilung bibliographischer Datenbanken angewendet werden. Dabei werden z.T. geringfügige Anpassungen einzelner Kriterien notwendig, die in der Charakteristik der SIGLE-Datenbank begründet sind. Im Rahmen der Beurteilung wird

¹⁴³ Eine vollständige Liste der Kriterien findet sich im Anhang der Arbeit. Vgl. dazu Kapitel VII.

¹⁴⁴ Armstrong zitiert nach Cooke et. al. (1996), S. 38.

¹⁴⁵ Vgl. Kapitel I Abschn. 1.1.

an den entsprechenden Stellen darauf hingewiesen.

2. Kriterien zur Beurteilung der CORDIS-Datenbanken

Die beiden CORDIS-Datenbanken Ergebnisse und Veröffentlichungen sollen mit Hilfe der Selection criteria for quality controlled information gateways beurteilt werden.

Dabei sollen mit Ausnahme der fünften Kategorie, der Collection Management Policy, alle Kategorien zur Anwendung kommen.

Die Collection Management Policy untersucht den Wert einer Internet-Ressource im Umfeld eines subject gateways. Sie bildet damit die Grundlage für die Verzeichnung oder Löschung eines Dokuments im konkreten gateway. Zur Prüfung der Qualität einer Datenbank ist sie jedoch nicht geeignet. Bei der Beurteilung der CORDIS-Datenbanken bleibt der Bereich der Collection Management Policy daher unberücksichtigt.

V. Die Anwendung der ausgewählten Kriterien zur Qualitätsbeurteilung auf die SIGLE- und CORDIS-Datenbanken

1. Beurteilung der SIGLE-Datenbank

Zur Qualitätsbeurteilung der SIGLE-Datenbank wurde zunächst die Aktualität (Timeliness) untersucht. Dazu mußte die Aktualität der Updates, bzw. die Aktualität des in einem Update enthaltenen Materials, festgestellt werden.¹⁴⁶ Anhand der vorhandenen „Summary Sheets“, aber auch im Rahmen von Datenbankrecherchen konnte festgestellt werden, daß die SIGLE-Datenbank monatlich aktualisiert wird. Die Updates erfolgen gegen Ende eines Monats, scheinbar aber zu keinem konkret festgelegten Termin. In den Bearbeitungszeitraum der vorliegenden Arbeit fielen somit drei Updates, das Datum des letzten Updates wird dem Benutzer beim Einwählen in die Datenbank angezeigt.

Zur Feststellung der Aktualität der in SIGLE verzeichneten Literatur, wurden stichprobenartig die Updates verschiedener Jahre näher untersucht. Untersuchungsgrundlage war dabei letztlich die Gesamtheit der innerhalb eines Jahres

¹⁴⁶ Bzgl. der einzelnen Kriterien vgl. die Liste der SCOUG im Anhang dieser Arbeit unter VII, oder die Beschreibung der Kriterien in Kapitel III. Abschn. 1.3.

eingespielten Datensätze.¹⁴⁷ Zur Vermeidung verzerrter Werte wurden die Updates früherer Jahre verwendet. Die Vorgehensweise entsprach dabei den von Jascó vorgeschlagenen Methoden.¹⁴⁸

Im folgenden soll beispielhaft die Untersuchung der Updates aus den Jahren 1990 und 1996 dargestellt werden.

In einem ersten Schritt wurden in der SIGLE-Datenbank insgesamt 42.937 Datensätze mit Erscheinungsjahr 1990 festgestellt. Danach mußten die einzelnen in SIGLE verzeichneten Dokumentarten durch eine Abfrage des entsprechenden Index-Feldes ermittelt werden. Für die insgesamt zehn verschiedenen Dokumenttypen, u.a. Reports, Conference Proceedings, Dissertationen und Progress Reports, wurden sodann alle Datensätze mit Erscheinungsjahr 1990 bestimmt. In einem weiteren Schritt wurde sowohl für die gesamte SIGLE-Datenbank als auch für die einzelnen Dokumentarten festgestellt, wieviel Prozent der einzelnen Datensätze mit Erscheinungsjahr 1990, bereits 1990 in die SIGLE-Datenbank eingespielt waren. Für die gesamte SIGLE-Datenbank ergab sich dabei eine Aktualitätsrate von ca. 13,8%.¹⁴⁹ Bei den einzelnen Dokumenttypen war die Aktualitätsrate sehr unterschiedlich. Abbildung 2 auf der nächsten Seite soll die Vorgehensweise erläutern.

¹⁴⁷ Um die monatlichen Updates untersuchen zu können, hätte das genaue Aktualisierungsdatum bekannt sein müssen. In SIGLE kann im Update-Feld nur nach dem genauen Datum oder einer Jahreszahl recherchiert werden.

¹⁴⁸ Vgl. Jascó (1992), S. 28-33.

¹⁴⁹ Der Wert ergibt sich bei einer Gegenüberstellung der 42.937 Datensätze mit Erscheinungsjahr 1990 und der 33.718 Datensätze, die 1990 in die SIGLE-Datenbank eingespielt wurden. Nur 4.640 Datensätze weisen beide Merkmale auf, das entspricht ca. 13,8%.

```

FILE 'SIGLE' ENTERED AT 11:08:55 ON 02 SEP 1999
FILE LAST UPDATED: 20 AUG 1999      <19990820/UP>
FILE COVERS 1980 TO DATE.

I) Ermittlung der Gesamtzahl der Dokumente mit Erscheinungsjahr 1990
=> s py=1990
L5      42937 PY=1990

II) Dokumenttypen mit Erscheinungsjahr 1990
Ermittlung der Gesamtzahl der verz. Reports
=> s r/dt
L11      190013 R/DT

Ermittlung der Reports aus 1990
=> s l5 and l11
L22      11132 L5 AND L11

III) Gesamtzahl der 1990 in die SIGLE-Datenbank eingespielten Datensätze
=> s 1990/up
L27      33718 1990/UP      (19900000-19909999/UP)_

IV) Bestimmung der Werte für einzelne Dokumenttypen
Eingespielte Reports in 1990 mit Erscheinungsjahr 1990
=> s l22 and l27
L29      3078 L22 AND L27_      Dies entspricht ca. 28%.

Eingespielte Progress Reports aus 1990
=> s l24 and l27      Dies entspricht ca. 20%.
L36      424 L24 AND L27_

Eingespielte Bibliographien aus 1990
=> s l16 and l27      Dies entspricht ca. 14%
L30      5 L16 AND L27_

Eingespielte Dissertationen aus 1990
=> s l18 and l27      Dies entspricht ca. 1%
L32      248 L18 AND L27_

```

Abbildung 2: Auszug aus einer SIGLE-Recherche¹⁵⁰

Bei Tagungsberichten (Conference Proceedings) waren nur 13% in 1990 eingespielt. Ein Wert von ca. 43% wurde im Bereich der Programm Documentation errechnet.

Für 1996 ergibt sich ein ähnliches Bild. Nur ca. 34% der Reports mit Erscheinungsjahr 1996 waren bereits im gleichen Jahr in der SIGLE-Datenbank verzeichnet. Für die Conference Proceedings war der ermittelte Wert mit nur 4% noch unbefriedigender. Die gesamte SIGLE-Datenbank weist 1996 einen etwas besseren Aktualitätsgrad als 1990 auf. Im Rahmen der Recherchen in der Datenbank wurde ein Wert von 25% festgestellt.

Insgesamt ergibt sich, daß die Dokumente erst nach mehreren Jahren komplett verzeichnet sind. So wurden z.B. im Jahr 1996 noch Datensätze mit Erscheinungsjahr 1990 in die SIGLE-Datenbank nachgemeldet. Dies ist nach Meinung der Verfasserin insbesondere bzgl. der verzeichneten Reports und Conference Proceedings ein zu langer Zeitraum. Gerade in diesen Bereichen ist die Aktualität sehr wichtig, da z.B.

¹⁵⁰ Die Prozentangaben wurden gerundet.

viele Tagungsberichte sehr schnell vergriffen und später zudem nicht mehr erhältlich sind. Dasselbe gilt für Forschungsberichte, die in der Regel sehr spezielle und aktuelle Ergebnisse im entsprechenden Fachgebiet wiedergeben und daher einer sehr schnellen Veralterung unterliegen. Sie müssen somit entsprechend schnell zur Verfügung gestellt, bzw. möglichst schnell nachgewiesen werden. Der geringe Grad an Aktualität ist in diesem Bereich auch deswegen besonders mißlich, da die SIGLE-Datenbank auf europäischer Ebene ein Hauptinstrument zur Ermittlung von Forschungsberichten darstellt.

Bezüglich der Konsistenz (**Consistency**), welche die SCOUG im Bereich der Feldanordnungen und –bezeichnungen sowie bei anderen Datenelemente und z.B. der Indexierungspolitik fordert, ergibt sich ein anderes Bild. Zu ihrer Prüfung wurden in der SIGLE-Datenbank Datensätze der verschiedenen Dokumenttypen ermittelt und sodann die Feldbezeichnungen und -anordnungen verglichen. Dabei zeigte sich, daß die in der Datenbank erfaßten Datensätze eine einheitliche Struktur aufweisen und daher konsistent sind. Abweichungen finden sich lediglich im Bereich der Abstracts und der Schlagwortvergabe, da nicht alle Datensätze solche Angaben aufweisen. Zu bemängeln ist allerdings, daß nicht festgestellt werden kann, in welchen Datensätzen keine Abstracts oder Schlagworte enthalten sind. Die Datenbankbeschreibung der SIGLE-Datenbank stellt nur fest: "*many records include also keywords and abstracts*".¹⁵¹ Auch durch eine Recherche in der Datenbank kann nicht festgestellt werden, welche bzw. wieviel Datensätze keine Abstracts oder Schlagworte erhalten haben. Dies ist darin begründet, daß SIGLE eine Recherche nach Abstracts oder auch eine vollständige Trunkierung grundsätzlich nicht zuläßt. Weiterhin wurde geprüft, ob die Verzeichnung der unterschiedlichen Dokumenttypen über mehrere Jahre hinweg konstant geblieben ist (Konsistenz der Indexierung). Dazu wurde beispielhaft für die Dokumenttypen Reports, Conference Proceedings, Progress Reports und Dissertationen die Gesamtzahl der verzeichneten Datensätze für verschiedene Jahre ermittelt. Abbildung 3 soll dies verdeutlichen.

¹⁵¹ SIGLE Description (1999) unter <http://www.FIZ-Karlsruhe.DE/stn/Databases/sigle.html>.


```
FILE 'SIGLE' ENTERED AT 09:42:22 ON 08 SEP 1999
COPYRIGHT 1999 (C) EUROPEAN ASSOCIATION FOR GREY LITERATURE
EAGLE)
```

Erscheinungsjahr 1993

Reports

=> _s 14 and 18

L24 11925 L4 AND L8_

Progress Reports

=> _s 14 and 19

L25 2273 L4 AND L9_

Conference Proceedings

=> _s 14 and 110

L26 787 L4 AND L10_

Dissertationen

=> _s 14 and 111

L27 9413 L4 AND L11_

Erscheinungsjahr 1994

Reports

=> _s 15 and 18

L28 10549 L5 AND L8_

Progress Reports

=> _s 15 and 19

L29 1881 L5 AND L9_

Conference Proceedings

=> _s 15 and 110

L30 872 L5 AND L10_

Dissertationen

=> _s 15 and 111

L31 9497 L5 AND L11_

Abbildung 3: SIGLE-Recherche bzgl. Konsistenz der Indexierung.

Ähnliche Ergebnisse ergeben sich auch für weitere Jahre. Obwohl die Anzahl der Datensätze einzelner Dokumenttypen z.B. bei Dissertationen oder Reports gelegentlich abweicht, bleibt das Verhältnis der Dokumenttypen zueinander über einen längeren Zeitraum praktisch konstant, eine Konsistenz der SIGLE-Datenbank bzgl. der Indexierung ist daher vorhanden.

Im Bereich von **Coverage / Scope** war zu untersuchen, wie gut die SIGLE-Datenbank ihren Themenbereich bzw. ihr Fachgebiet erfaßt. Dazu wurde geprüft, inwieweit die in den Dokumentations- und Hilfsmaterialien verzeichneten Themen und Dokumentarten auch tatsächlich in der Datenbank verzeichnet sind. Den Angaben dieser Materialien zufolge, verzeichnet die Datenbank Dokumente aus insgesamt 16 Themenbereichen, u.a. aus "Aeronautics and Space Technology", "Agriculture, Plants and Veterinary Sciences", "Humanities, Psychology and Social Sciences", "Biology and Medicine" sowie "Energy and Power", "Physics and Mathematical Sciences" und "Mechanical, Industrial, Civil and Marine Engineering".¹⁵² Alle Themenbereiche lassen sich bei

¹⁵² Vgl. SIGLE-Summary Sheet (1999) unter <http://info.cas.org/ONLINE/DBSS/sigless.html>.

einer Recherche in der Datenbank nachweisen, wenn z.T. auch in sehr geringer Konzentration. Im Bereich von "Aeronautics and Space Technology" z.B., finden sich insgesamt nur 234 Datensätze, bei "Computer Science" sogar nur 16 und bei "Missile Technology" lediglich 25 Nachweise.

Die in den Dokumentations- und Hilfsmaterialien angegebenen Dokumenttypen¹⁵³ können in der Datenbank nachgewiesen werden. Bedauerlicherweise aber nicht in der im STN Newsletter veröffentlichten Konzentration.¹⁵⁴ Die Datensätze der SIGLE-Datenbank bestehen nicht zu 55% aus Hinweisen auf Reportliteratur; selbst bei einer Zusammenfassung, der durch eine Recherche ermittelten Werte für Reports und Progress Reports, lassen sich nur 34% nachweisen.¹⁵⁵

```
SIGLE-Recherche vom 02.09.1999
Download Erstellungsdatum 02-09-1999
Download Erstellungszeit 10:46:41

Ermittlung der Gesamtzahl der in SIGLE verzeichneten Dokumente

Recherche im Update-Feld. Da der Berichtszeitraum lt. SIGLE-Dokumentation
1980 ff. umfaßt Recherche mit:
=> s up>1980
L1      629545 UP>1980
        (UP>19809999)

Ermittlung der Gesamtanzahl an Reports
=> s r/dt
L8      190013 R/DT      Beide Dokumenttypen zusammengerechnet er-
                        geben 210.629 Datensätze, das entspricht ca. 34%.

Ermittlung der Gesamtzahl der Progress Reports
=> s pr/dt
L10     20616 PR/DT
```

Abbildung 4: Auszug aus einer SIGLE-Recherche

Dies ist um so erstaunlicher, da im Juli 1999 die Datenbank FTN (Forschungsberichte aus Technik und Naturwissenschaften) in die SIGLE-Datenbank integriert wurde und die Verfasserin eigentlich erwartet hätte, daß der angegebene Anteil an Reportliteratur noch ansteigt.

Lücken bezüglich der versprochenen Themen und Dokumentarten wurden in der SIGLE-Datenbank insgesamt nicht festgestellt.¹⁵⁶ Die Verfasserin ist allerdings der Ansicht, daß die Nutzer stärker als bisher auf die Schwerpunkte der Datenbank im Bereich der Naturwissenschaften hingewiesen werden sollten. Im Bereich von Coverage und Scope sollte zusätzlich erwähnt werden, daß die SIGLE-Datenbank

¹⁵³ Vgl. SIGLE-Summary Sheet (1999) unter <http://info.cas.org/ONLINE/DBSS/sigless.html>.

¹⁵⁴ STN Newsletter (April 1999).

¹⁵⁵ Vgl. Abbildung 4.

¹⁵⁶ Bzgl. der geforderten Informationen über Neuerungen und Änderungen, vgl. den folgenden

auch Datensätze mit einem Erscheinungsjahr vor 1980 verzeichnet, obwohl ihr Berichtszeitraum nach den Angaben im Summary Sheet im Jahr 1980 beginnt. Der Anteil dieser Datensätze ist mit ca. 1,19% allerdings eher gering.

In einem nächsten Schritt waren die zur SIGLE-Datenbank erhältlichen Hintergrund- und Hilfsmaterialien zu untersuchen (**Documentation**). Nach dem SCOUG-Modell ist in diesem Bereich insbesondere darauf zu achten, ob die vorhandenen Materialien aktuell, genau, verständlich sind. Darüber hinaus sollten die Hilfs- und Dokumentationsmaterialien kostenfrei oder gegen ein geringes Entgelt zu beziehen sein.

Für die SIGLE-Datenbank sind sowohl gedruckte als auch Online-Hilfen erhältlich. Die neueste Auflage des zur Recherche unabdingbaren Summary Sheets weist einen Stand von Juli 1999 auf¹⁵⁷, die Aktualität des Hilfsmaterials ist somit gegeben. Die Abfrage der Summary Sheets über das Internet ist kostenfrei; weitere Hilfsmittel, wie z.B. der STNGUIDE können ebenfalls kostenfrei beim FIZ Karlsruhe bestellt werden. Die für die SIGLE-Datenbank vorhandene Online-Hilfe ist in verschiedene Abschnitte gegliedert und enthält ebenso wie die Print-Ausgabe Angaben zu suchbaren Feldern, Ausgabeformaten, Operatoren, usw. Eine Übersicht gibt Abbildung 5.

```
FILE 'SIGLE' ENTERED AT 09:50:49 ON 06 SEP 1999
COPYRIGHT 1999 (C) EUROPEAN ASSOCIATION FOR GREY LITERATURE (EAGLE)

The following HELP messages are available to obtain information
on the SIGLE file:

HELP ACCESSION - SIGLE accession number formats
HELP CONTENT   - general SIGLE file description
HELP COST      - price schedule for the SIGLE file
HELP CROSSOVER - file crossover searching in SIGLE
HELP DESK      - information on SIGLE file user assistance
HELP DFIELDS   - list of display field codes
HELP EFIELDS   - list of select fields
HELP FIELDS    - list of field and format help messages for the
                  SIGLE file
HELP FORMAT    - predefined formats for display and print
HELP HIGHLIGHT - highlighting information for the SIGLE file
HELP (L)       - (L) operator use in the SIGLE file
HELP RANGE     - RANGE parameters for the SIGLE file
HELP (S)       - (S) operator use in the SIGLE file
HELP SFIELDS   - list of search field codes
HELP SRTFIELDS - list of sortable fields in the SIGLE file
HELP UPDATE/SDI - manual and automatic update searching
HELP USAGETERMS - use and distribution restrictions applicable
                  to the SIGLE file
```

Abbildung 5: Verfügbare Online-Hilfen in der SIGLE-Datenbank

Im Rahmen der Online-Hilfen ist zu beachten, daß nicht für jedes Feld der SIGLE-

Datenbank spezifische Helfefunktionen verfügbar sind. Die Fehlermeldungen wiederum sind alle auf konkrete, fehlerhafte Eingaben bezogen, Kontext-Sensitivität ist insofern gegeben. Neben den gedruckten Materialien sind Hilfen für die SIGLE-Datenbank auch über ein Help Desk beim FIZ-Karlsruhe und eine Telefon-Hotline zu erhalten. Über die Hotline kann spezifischer Support angefordert werden. Zeitweise Störungen bzw. Zugangsprobleme zu den Datenbanken, werden bei STN durch die sogenannte "Alert Line", einen automatischen telefonischen Ansagedienst, bekanntgegeben. Neuerungen und Änderungen bzgl. der Datenbank können sowohl in der Datenbank selbst als auch über den regelmäßig erscheinenden STN Newsletter abgefragt werden.

Die Verfasserin ist der Ansicht, daß für die SIGLE-Datenbank insgesamt ein sehr umfangreiches System an Hilfen und Dokumentationsmaterialien verfügbar ist. Bezüglich der vorhandenen Dokumentation ist festzustellen, daß die Materialien zwar keine grundsätzlich falschen Angaben beinhalten, jedoch gelegentlich Mängel im Bereich der Genauigkeit aufweisen. Diese Mängel wurden in den Abschnitten zur „Consistency“ und „Coverage and Scope“ bereits beschrieben. Aus diesem Grund wird hier nicht näher auf sie eingegangen. Die Benutzbarkeit der gedruckten Hilfs- und Dokumentationsmaterialien könnte zudem verbessert werden, indem z.B. die Angaben über die Erreichbarkeit der Hotline und der Alert Line oder die Öffnungszeiten des Help Desks, obwohl sie nicht nur für die SIGLE-Datenbank gelten, in das entsprechende Summary Sheet integriert würden. Nutzern der SIGLE-Datenbank bliebe so ein lästiges Zusammentragen der einzelnen Informationen erspart.

Ein weitere Kategorie des SCOUG-Modells, mit der die SIGLE-Datenbank bewertet werden soll, ist die Genauigkeit und Fehlerrate der Datenbank (**Accuracy/Error rate**). Bzgl. der Genauigkeit sind die der Datenbank zugrunde liegenden Quellen zu prüfen. Die einzelnen Dokumenttypen der SIGLE-Datenbank sowie deren Aktualität sind bereits im Rahmen der "Timeliness"-Beurteilung untersucht worden. An dieser Stelle hätte daher noch die Vollständigkeit der Quellen geprüft werden müssen. Dies erwies sich jedoch als äußerst schwierig, da in der SIGLE-Datenbank, keine vollständige Trunkierung zugelassen ist. Im Rahmen der stichprobenartig durchgeführten Recherchen konnte lediglich festgestellt werden, daß bei insgesamt 11.446 Datensätzen (ca. 2%) kein Erscheinungsjahr erfaßt ist. Außerdem wurden bei einer Durchsicht des entsprechenden Index-Feldes 124

¹⁵⁷ Abfragbar über <http://info.cas.org/ONLINE/DBSS/sigless.html>.

Datensätze ohne Sprachbezeichnung ermittelt.¹⁵⁸ Weitere Angaben zur Vollständigkeit, z.B. bzgl. der Dokumenttypen, konnten bedauerlicherweise nicht ermittelt werden.

¹⁵⁸ Im Index-Feld wird durch den Eintrag „unavailable“ auf die fehlende Sprachbezeichnung hingewiesen.

Zur weiteren Beurteilung der Genauigkeit und der Fehlerrate wurde nach Schreibfehlern in der Datenbank, insbesondere in wichtigen Feldern, wie z.B. im Autoren-Feld, gesucht. Dazu wurden, wie von Jascó vorgeschlagen, verschiedene Indices getestet.¹⁵⁹ Im Rahmen dieser Recherchen wurden keine Schreibfehler gefunden, von einer Genauigkeit der Daten war daher auszugehen. Diese Genauigkeit der Daten wird in der SIGLE-Datenbank durch festgelegte formale Anforderungen bei der Datenerfassung und mehrmaliges Prüfen der Datensätze vor dem Einspielen in die Datenbank erreicht.¹⁶⁰ Sollten dennoch im Rahmen einer SIGLE-Recherche Fehler festgestellt werden, können Nutzer diese Fehler melden, müssen dazu aber die Datenbank verlassen und z.B. über STNMAIL eine Email an das Help Desk des FIZ Karlsruhe richten.¹⁶¹ Eine Möglichkeit, fehlerhafte Datensätze, ähnlich dem von Mintz vorgeschlagen "FIXIT"-Kommando¹⁶², in der Datenbank zu markieren, besteht in der SIGLE-Datenbank nicht.

In einem weiteren Schritt wurden die Zugänglichkeit und Benutzbarkeit der SIGLE-Datenbank (**Accessibility/Ease of use**) untersucht. Dabei waren sowohl der Zugang zum System als auch der Zugang zu den Daten zu beurteilen. Entsprechend den Kriterien der SCOUG mußte u.a. auf das Vorhandensein von Proximity-Operatoren sowie die Möglichkeit einer Phrasenrecherche geachtet werden. Weiterhin waren z.B. die Suchbarkeit der verschiedenen Felder sowie die Tiefe und Flexibilität der Indexierung zu prüfen (Zugang zu den Daten).

Zunächst soll auf den Zugang zum Datenbanksystem von STN und zur SIGLE-Datenbank eingegangen werden. Recherchen in der SIGLE-Datenbank bzw. in den Datenbanken von STN können mit einem Standard-PC durchgeführt werden. Empfehlungen bzgl. der Ausstattung (Größe der Festplatte oder des Arbeitsspeichers) werden potentiellen Nutzern durch das FIZ Karlsruhe, den deutschen Vertragspartner von STN, gegeben.¹⁶³ Positiv ist in diesem Zusammenhang zu bewerten, daß der zur Recherche erforderliche Nutzungsvertrag auch Online abgeschlossen werden kann.¹⁶⁴ Der Zugang zum Server von STN und damit auch zu SIGLE ist damit sofort möglich. Erforderliche Informationen können so ohne Zeitverlust ermittelt werden. Der Zugang zu den Datenbanken kann

¹⁵⁹ Vgl. Jascó (1993 III), S. 30-36.

¹⁶⁰ Iser, Uschi <ui@fiz-karlsruhe.de> Persönliche Email (14.09.1999).

¹⁶¹ Vgl. FIZ/STN Global Hotline (1998) unter <http://www.fiz-karlsruhe.de/stn.html>.

¹⁶² Vgl. Mintz (1990), S. 17.

¹⁶³ Beschreibung der erforderlichen Hardware-Komponenten (1997) unter <http://www.fiz-karlsruhe.de/fiz/faq/hardware.html>.

¹⁶⁴ Nutzungsbedingungen mit anschließendem Vertragsformular (1999) unter <http://stnweb.fiz-karlsruhe.de>.

bei STN sowohl direkt als auch über einen gateway erfolgen. Einzelne Zugangsmöglichkeiten können beim FIZ Karlsruhe erfragt werden.¹⁶⁵

Im folgenden sollen die Zugangsmöglichkeiten zu den eigentlichen Daten der SIGLE-Datenbank beurteilt werden. STN unterstützt die von der SCOUG geforderte Recherche mit Proximity-Operatoren; auch eine Phrasensuche ist möglich. Nicht vorhanden ist dagegen eine automatische Suche nach Synonymen sowie der Pluralform von Suchbegriffen. Die Recherche-Unterstützung in Form der automatischen Synonymabfrage wäre vor allem dann hilfreich, wenn der Rechercheur kein Spezialist im entsprechenden Fachgebiet der Datenbank ist. Der Zugang zu den Daten erfolgt in der SIGLE-Datenbank grundsätzlich über die Retrievalsprache Messenger, auf eine weitere Recherchemöglichkeit über das System STN Easy wird an späterer Stelle eingegangen¹⁶⁶. Eine Recherche kann in insgesamt 17 verschiedenen Feldern erfolgen. Die Suchbegriffe können dabei keinem Online-Thesaurus entnommen werden.¹⁶⁷

```
=> _help sfields
```

| Search Code | Definition |
|-------------|-------------------------------|
| /AN | Accession Number |
| /AU | Author |
| /BI | Basic Index |
| /CC (/CCEN) | Classification Code |
| /CCDE | Classification Code in German |
| /CCFR | Classification Code in French |
| /CS | Corporate Source |
| /CY | Country of Publication |
| /DT (/TC) | Document Type |
| /LA | Language |
| /NC | Number of Contract |
| /NR | Number of Report |
| /PD | Publication Date |
| /PY | Publication Year |
| /SO | Source |
| /ST | Supplementary Term |
| /TI | Title |

Abbildung 6: Recherchierbare Felder in der SIGLE-Datenbank.

Welche Felder im einzelnen suchbar, welche nur anzeigbar sind kann also über die Online-Hilfe abgefragt werden. Bzgl. der Flexibilität und Tiefe der Indexierung ist festzuhalten, daß die Datensätze der SIGLE-Datenbank durch Schlagworte und die Classification Codes erschlossen werden. Es können bis zu zwölf dieser Schlagworte

¹⁶⁵ Abfrage über <http://www.fiz-karlsruhe.de/stn.html> (Stand 1998).

¹⁶⁶ Vgl. Abschnitt 1.2 in diesem Kapitel.

¹⁶⁷ Vgl. Abbildung 6.

vergeben sein, denen kein bestimmtes Regelwerk zugrunde zu liegen scheint. Falls Schlagworte vergeben sind¹⁶⁸, zeichnet sich die SIGLE-Datenbank nach Meinung der Verfasserin durch eine flexible und tiefe Indexierung aus. Insgesamt liegt eine gute Zugänglichkeit und Bedienbarkeit (Accessibility/Ease of use) der SIGLE-Datenbank vor. Sowohl der Zugang zum System als auch zu den eigentlichen Datensätzen funktionierte während des gesamten Bearbeitungszeitraums ordnungsgemäß.

Neben den soeben beschriebenen fünf Kriterien des "Grundmodells" der SCOUG sollen im folgenden die Kriterien zur Beurteilung bibliographischer Datenbanken angewendet sowie die Endnuterorientierung der SIGLE-Datenbank bewertet werden.

Zunächst soll untersucht werden, inwieweit die SIGLE-Datenbank den Anforderungen der SCOUG an bibliographische Datenbanken entspricht. In der SIGLE-Datenbank kann die Anzahl erzielter Rechercheergebnisse durch eine Beschränkung auf bestimmte Fachgebiete und Sprachen reduziert werden. Die Quellen-Angaben im entsprechenden Feld sind ausgeschrieben und alle Felder mit Ausnahme des Abstract-Feldes suchbar. Die SIGLE-Datenbank entspricht der üblichen Struktur bibliographischer Datenbanken, sie enthält keine Wertung der zitierten Quellen. Insofern erfüllt sie daher die Anforderungen der SCOUG an eine bibliographische Datenbank. Im Gegensatz dazu, erfüllen die fehlende „author affiliation“ sowie die Tatsache, daß im Autoren Feld nur in der fest vorgegebenen Form (Name, Vorname/Au) recherchiert werden kann, nicht die Anforderungen der SCOUG. Die Integration von Erscheinungsdatum und -jahr im Quellen-Feld bei Anzeige der Suchergebnisse und die fehlende Verbindung zum Volltext der in SIGLE zitierten Literatur, sind in diesem Zusammenhang ebenfalls zu nennen. Die Verfasserin ist allerdings der Ansicht, daß die letzten Aspekte nicht unbedingt negativ zu bewerten sind. Anstelle der direkten Verbindung zum Volltext enthält die Datenbank Angaben zur Verfügbarkeit der entsprechenden Dokumente, so daß sie direkt, auch bei STN, bestellt werden könnten. Durch die Anzeige des Erscheinungsdatums oder -jahres im Quellen-Feld, wird die Übersichtlichkeit sogar verbessert, da alle zur Beschaffung des Dokuments erforderlichen Angaben an einer Stelle verzeichnet sind. Zusammenfassend ist festzustellen, daß die Anforderungen des SCOUG-Modells an bibliographische Datenbanken durch die SIGLE-Datenbank weitestgehend erfüllt werden.

¹⁶⁸ Vgl. dazu den Abschnitt „Consistency“.

Ein anderes Bild ergibt sich bei den Kriterien zur Endnutzerorientierung. In der SIGLE-Datenbank erhalten Endnutzer keine Hilfen bei der Formulierung von Suchanfragen. Die durch die SCOUG geforderten unterschiedlichen Abfrage-Levels sind ebenfalls nicht vorhanden. In SIGLE ist lediglich eine Recherche mit der Retrievalsprache Messenger vorgesehen. Zudem fehlt der *"live online customer support"*.¹⁶⁹ Das Help Desk ist zwar tagsüber erreichbar, aber nicht direkt über die SIGLE-Datenbank¹⁷⁰. Obwohl die eigentliche Abfrage eines Suchbegriffs in der SIGLE-Datenbank kostenfrei ist, entspricht auch die Kostenstruktur der Datenbank nicht den Anforderungen der SCOUG an Endnutzersysteme, da sowohl Gebühren für die Anschaltzeit als auch zusätzliche Kosten für die Anzeige von Dokumenten zu entrichten sind. Aus dem soeben gesagten ergibt sich, daß die SIGLE-Datenbank den Anforderungen der SCOUG an endnutzerorientierte Datenbanken nicht entspricht. Dies könnte aber auch darin begründet sein, daß Endnutzern eine zweite Recherchemöglichkeit zur Verfügung steht, auf die im folgenden eingegangen werden soll.

2. Die Recherche in der SIGLE-Datenbank mit STN Easy

In der SIGLE-Datenbank kann auch im Rahmen eines STN Easy-Zugangs recherchiert werden. STN Easy ist ein maskenbasierter *"Point-and-Click-Zugang"*¹⁷¹ zu den Datenbanken von STN. Aufgrund der maskenbasierten Recherche, könnte diese Zugangsform für Endnutzer geeigneter sein, als die Online-Version der Datenbank. Ob dies tatsächlich der Fall ist, soll im folgenden untersucht werden.

In der STN Easy-Version der SIGLE-Datenbank stehen dem Nutzer grundsätzlich zwei verschiedene Suchmöglichkeiten zur Verfügung, die sogenannte "Einfache Suche" und die "Fortgeschrittene Suche". Ebenso wie im Rahmen des Online-Zugangs ist eine Phrasensuche und eine Suche mit Booleschen Operatoren möglich. Ein Erlernen der Retrievalsprache Messenger vor Recherchebeginn, ist nicht mehr notwendig. Die STN Easy-Version weist also auf den ersten Blick tatsächlich eine größere Endnutzerorientierung auf.

¹⁶⁹ Basch (1990 II), S. 23.

¹⁷⁰ Vgl. den Abschnitt „Accuracy / Error rate“ in diesem Kapitel.

¹⁷¹ STN Easy-Dokumentation unter <http://stneasy.fiz-karlsruhe.de/html/deutsch/help.html#about>.

Im Rahmen der durchgeführten Recherchen¹⁷² zeigten sich jedoch einige Besonderheiten, die unbedingt beachtet werden sollten. Zunächst sollte festgestellt werden, ob die über STN Easy zugängliche Fassung der SIGLE-Datenbank mit ihrer Online-Version übereinstimmt. Dies scheint im Bereich von „Coverage and Scope“ der Fall zu sein. Diesbezügliche Recherchen nach zuvor in der Online-Version ermittelten Datensätzen, ergaben gleiche Suchergebnisse. Darüber hinaus ergaben thematische Recherchen in der Regel übereinstimmende Treffermengen. Bedauerlicherweise ist es aufgrund der Maskenstruktur von STN Easy nicht möglich, völlig identische Suchanfragen an beide Versionen der Datenbank zu stellen, so daß keine genaueren Aussagen erfolgen können.

Die voreingestellten Masken verhindern zudem eine Prüfung der Aktualität und Konsistenz der SIGLE-Datenbank im STN Easy-Zugang. Zum einen kann die Gesamtzahl an Datensätzen nicht festgestellt werden, zum anderen sind von der Online-Version abweichende Dokumenttypen voreingestellt. Tagungsberichte, Reports und Progress Reports sind überhaupt nicht, dafür zusätzlich Bücher, Zeitschriften und der Dokumenttyp "Alle" eingestellt. Eine Änderung oder Anpassung der Masken ist nicht möglich, nur vorgegebene Einstellungen können gewählt werden. Abbildung 7 soll dies verdeutlichen.

¹⁷² Der zur Recherche erforderliche Nutzungsvertrag kann auch bei STN Easy online unter <http://stnweb.fiz-karlsruhe.de> (1999) geschlossen werden.

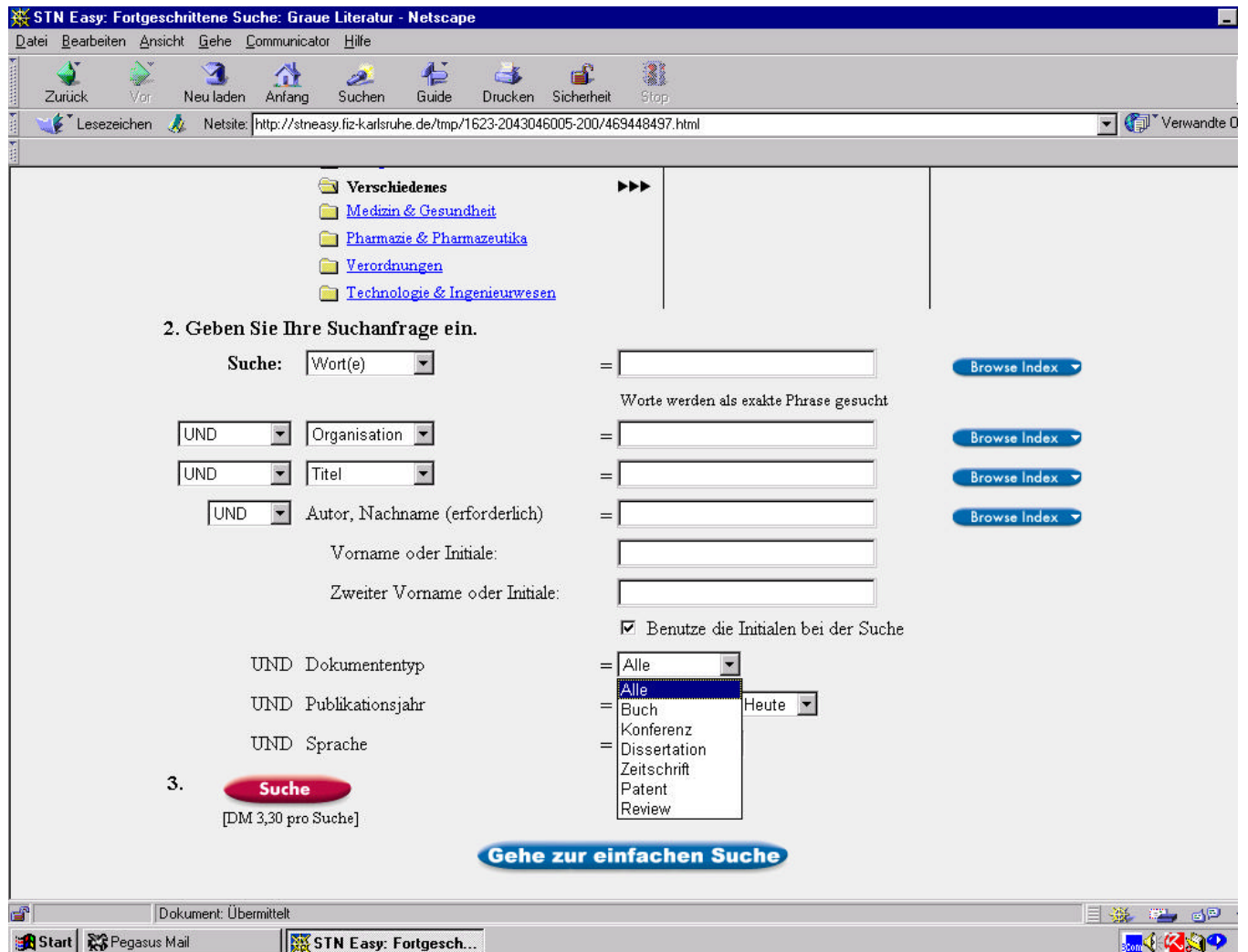


Abbildung 7: STN Easy-Suchmaske

Dies dürfte darin begründet sein, daß die Maskenstruktur für alle über STN Easy recherchierbare Datenbanken Gültigkeit haben soll. Eine Recherche mit aus der Online-Version bekannten Datensätzen ergab, daß auch in der STN Easy-Version der SIGLE-Datenbank Reports verzeichnet sind. Reports, Progress Reports und Tagungsberichte dürften daher implizit im Dokumententyp "Alle" enthalten sein. Reports, Tagungsberichte oder Progress Reports müßten im Rahmen einer Recherche demzufolge manuell aussortiert werden, da aus den Antwortübersichten nicht ersichtlich ist, um welchen Dokumenttyp es sich handelt. Je nach Größe der Suchergebnisse ist ein manuelles Sortieren der Ergebnisse äußerst zeit- und kostenintensiv. Bei Zugang über STN Easy ist in der SIGLE-Datenbank pro angezeigtem Datensatz eine Gebühr von DM 4,80 zu entrichten.¹⁷³ Zu bemängeln ist außerdem die Benutzbarkeit der Indices, die nur bei Eingabe eines Suchbegriffs im entsprechenden Feld aufgeblättert werden.

¹⁷³ Die Preise werden bei STN Easy in der entsprechenden Datenbank angezeigt.

Besonders mißlich ist dabei, daß immer nur ein kleiner Bereich des entsprechenden Indexes angezeigt wird und ein Vorwärts- oder Rückwärtsblättern nur dann möglich ist, wenn der Index zurückgesetzt und der Suchbegriff neu eingegeben wird.

Zusammenfassend läßt sich feststellen, daß die Notwendigkeit des manuellen Einschränkens ermittelter Suchergebnisse bzgl. der genannten Dokumenttypen in jedem Fall unzumutbar ist und aus Sicht der Endnutzerorientierung einen entscheidenden Qualitätsmangel darstellt. Dies auch insbesondere deswegen, weil die SIGLE-Datenbank, zumindest im Bereich der Forschungsberichte, ein Hauptinformationsinstrument darstellt. Endnutzer, welche aufgrund mangelnder Kenntnisse der Retrievalsprache Messenger in der STN Easy-Version der SIGLE-Datenbank recherchieren, müssen im Vergleich zur Online-Version, trotz wegfallender Anschlußgebühren¹⁷⁴ möglicherweise sehr viel höhere Kosten zahlen und wären dann benachteiligt. Die Verfasserin ist daher der Ansicht, daß die durch die SCOUG geforderte Endnutzerunterstützung auch bei einer Recherche über STN Easy nicht gegeben ist.

3. Beurteilung der CORDIS-Datenbanken

Im folgenden soll die Qualität der beiden CORDIS-Datenbanken *Ergebnisse* und *Veröffentlichungen* bewertet werden. Beide Datenbanken sollen dabei grundsätzlich gemeinsam untersucht werden; auf datenbankspezifische Besonderheiten wird an den entsprechenden Stellen gesondert hingewiesen.

Zunächst soll auf den Bereich der **Scope Policy** eingegangen werden. Die Selection criteria fordern an dieser Stelle eine Berücksichtigung der Informationsbedürfnisse potentieller Nutzer bei der Auswahl entsprechender Internet-Dokumente. In einem ersten Schritt war daher zu prüfen, inwieweit die beiden CORDIS-Datenbanken bei der Auswahl ihrer Informationen den Anforderungen ihrer Zielgruppe gerecht werden. Die Hauptzielgruppe des CORDIS-Systems sind europäische Unternehmen, insbesondere KMU (kleine und mittelständische Unternehmen), deren Wettbewerbsfähigkeit durch CORDIS verbessert werden soll.¹⁷⁵ Im einzelnen mußten die verzeichneten Themengebiete (Information Coverage), die Zugangsmöglichkeiten zu den Inhalten der Datenbanken (Access),

¹⁷⁴ Im Rahmen von STN Easy müssen keine Anschlußzeiten, sondern Gebühren von DM 3,30 pro durchgeführtem Suchvorgang gezahlt werden.

¹⁷⁵ Vgl. CORDIS-Einführung (1999) unter http://www.cordis.lu/de/src/s_016_de.htm.

aber auch die Breite und Tiefe der Indexierung (Cataloguing Policy , Geographical Issues) untersucht werden.

Die Datenbank *Ergebnisse* verzeichnet Ergebnisse aus Forschung und Entwicklung in den Bereichen Medizin, Wissenschaft und Technologie ungeachtet ihrer Finanzierungsquelle. Grundlage (Quelle) der verzeichneten Daten sind Meldungen forschender Organisationen sowohl aus dem öffentlichen als auch dem privaten Sektor, welche eine Verwertung ihrer Forschungsergebnisse anstreben.¹⁷⁶ Dazu sind in der Datenbank *Ergebnisse* die entsprechenden Organisationen mit ihren Ansprechpartnern erfaßt. Die dort verzeichneten Informationen stammen daher quasi von der Zielgruppe der Datenbank und sind demzufolge in jedem Fall angemessen. Der Schwierigkeitsgrad ist somit ebenfalls sachgemäß.

Die Datenbank *Veröffentlichungen* verzeichnet Hinweise, bibliographische Angaben sowie Auszüge aus Veröffentlichungen zu EG-finanzierten Forschungs- und Entwicklungsprogrammen.¹⁷⁷ Erfaßt sind Berichte (Forschungsberichte, Tagungsberichte, wissenschaftliche und technische Studien), aber auch Zeitschriftenartikel. Mit Hilfe der Datenbank *Veröffentlichungen* können sich Unternehmen über die Art und die Verfügbarkeit von Publikationen zu Gemeinschaftsprogrammen informieren. Die Angaben zur Verfügbarkeit sind dabei so detailliert, daß sie direkt zur Grundlage eines Bestellvorgangs gemacht werden können. Die Datenbank *Veröffentlichungen* dürfte daher dem Bedürfnis vieler Unternehmen nach schneller und gezielter Information entsprechen. Da die verzeichneten Quellen (Forschungsberichte, Tagungsberichte, Zeitschriftenartikel, usw.) den Informationsmitteln entsprechen, die bei einer Suche nach Informationen üblicherweise herangezogen werden, sind sie in jedem Falle, auch bezogen auf den Schwierigkeitsgrad, angemessen.

Der Zugang zu ihren Informationen ist in beiden CORDIS-Datenbanken maskenbasiert und kostenfrei. Nutzern steht eine Vielzahl suchbarer Felder zur Verfügung. In beiden Datenbanken können die Suchbegriffe einer Recherche frei eingegeben, oder bei bestimmten Feldern aus umfangreichen Indices übernommen werden.¹⁷⁸

¹⁷⁶ Datenbank-Beschreibung (1998) unter http://www.cordis.lu/de/src/a_018_de.htm.

¹⁷⁷ Datenbank-Beschreibung (1998) http://www.cordis.lu/de/src/a_017_de.htm.

¹⁷⁸ Vgl. dazu die Abbildungen 8 und 9.

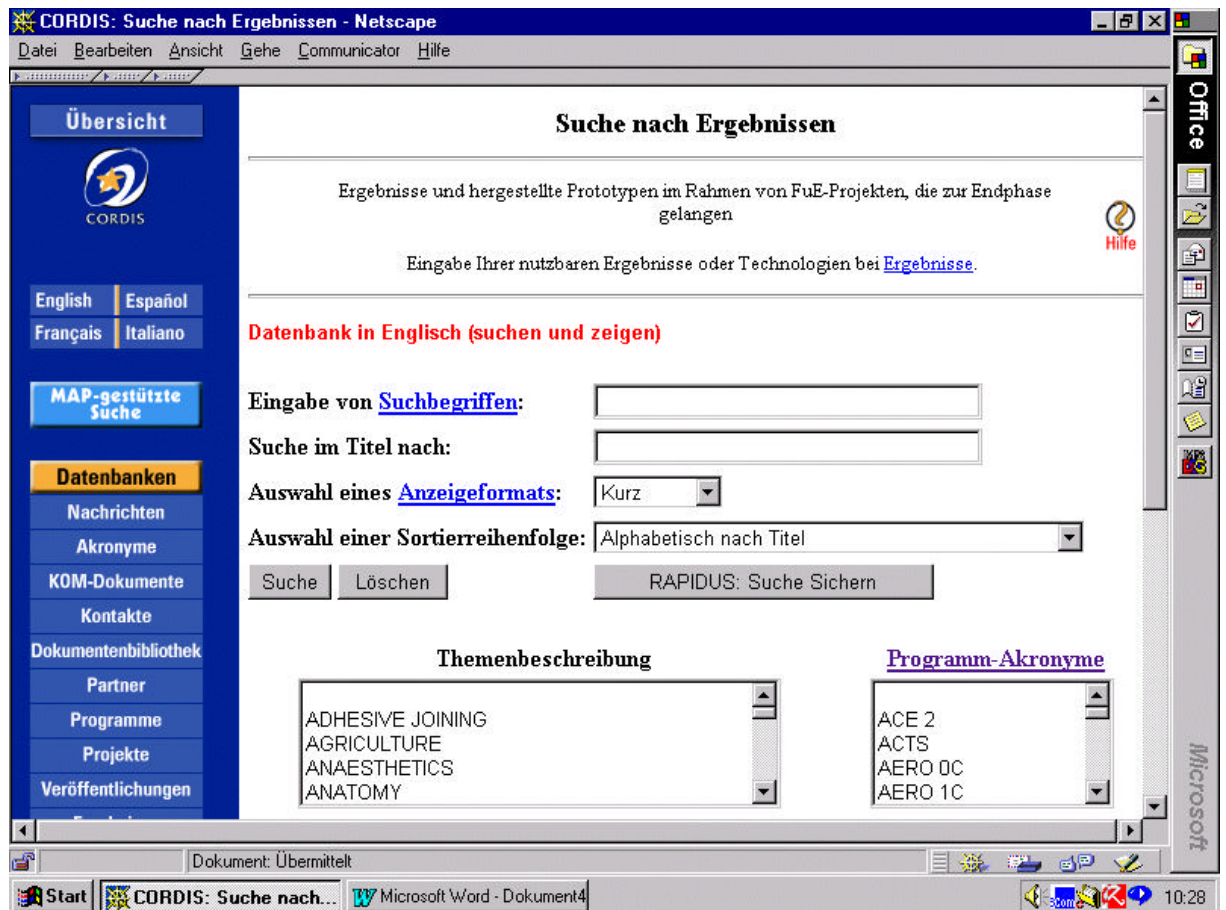


Abbildung 8: Datenbank *Ergebnisse*

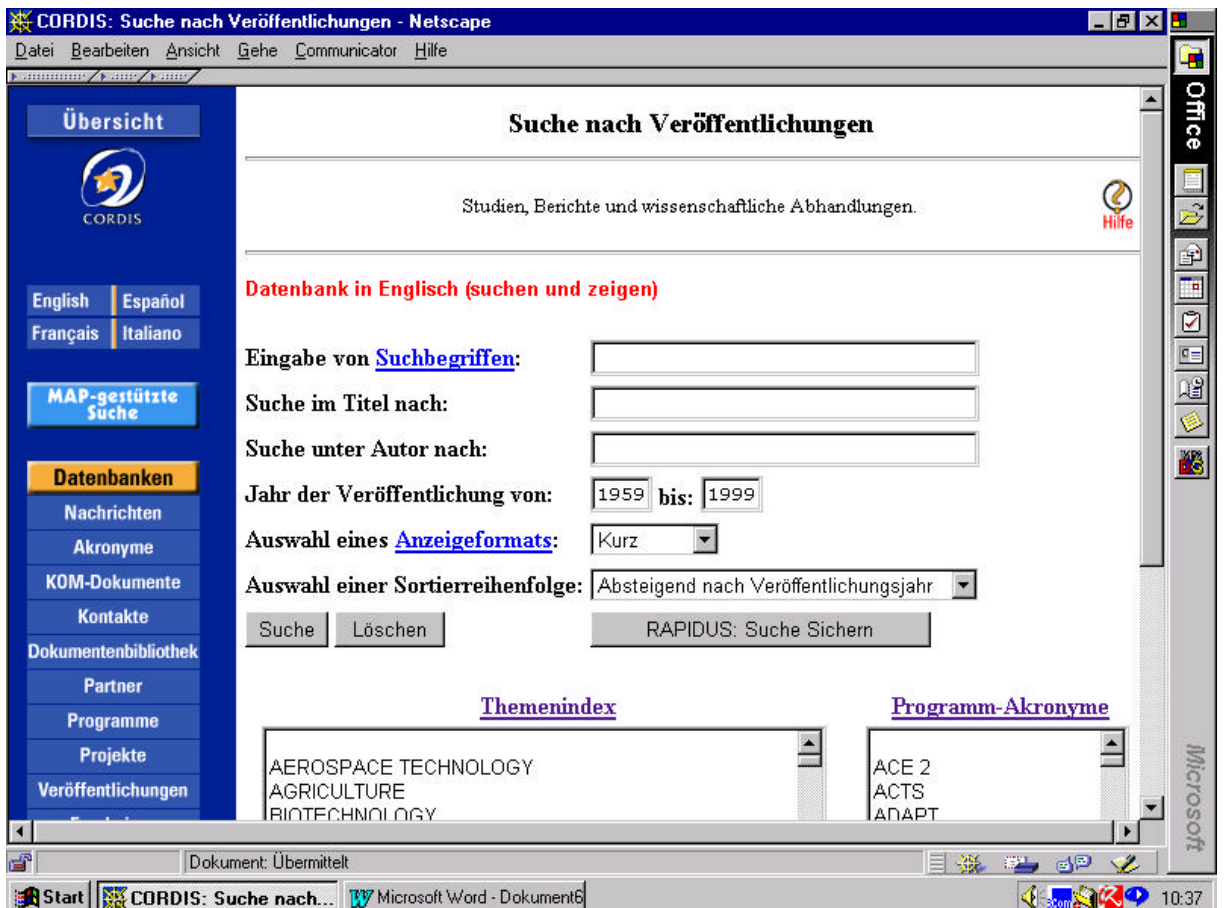


Abbildung 9: Datenbank *Veröffentlichungen*

Zur Nutzung beider CORDIS-Datenbanken, ist eine Registrierung grundsätzlich nicht erforderlich¹⁷⁹. Sie wird nur dann notwendig, wenn in den Datenbanken wiederholt nach den gleichen Themen recherchiert und daher ein Suchprofil (ähnlich den von Online-Datenbanken bekannten SDI-Diensten) abgespeichert werden soll. Für beide Datenbanken wird Standard-HTML verwendet¹⁸⁰, die Antwortzeiten beider Datenbanken sind sehr kurz, sie werden auch nicht durch überflüssige Frames oder Graphiken verlängert. Die verwendeten Technologien beeinträchtigen die Recherchierbarkeit der Datenbanken nicht, so daß von einer Angemessenheit ausgegangen werden dürfte. Beide CORDIS-Datenbanken sind nicht verschlüsselt und bieten somit keinen Sicherheitsschutz. Auf die mangelnde Sicherheit werden die Nutzer in beiden Datenbanken nur hingewiesen, wenn sie die sogenannten Seiteninformationen abrufen¹⁸¹. Dennoch ist insgesamt davon auszugehen, daß auch die Zugangsmöglichkeiten der Datenbanken *Ergebnisse* und *Veröffentlichungen* an die Bedürfnisse ihrer Nutzer angepaßt sind.

Als letzter Bereich der Scope Policy sollen die Breite und Tiefe der Indexierung der beiden Datenbanken untersucht (Cataloguing Policy), der geographische Raum aus dem die Forschungsergebnisse und Dokumente stammen sowie die Sprache der verzeichneten Dokumente festgestellt werden (Geographical Issues).

Die formale Erschließung erfolgt durch eine Vielzahl verschiedener, in beiden Datenbanken aber entsprechend ihrem jeweiligen Zweck unterschiedlicher Kategorien. In der Datenbank *Veröffentlichungen* stehen deshalb die bibliographischen Angaben (z.B. Autoren- und Titel-Feld) und die Angaben zur Verfügbarkeit der verzeichneten Dokumente (EUR-OP-Referenz-Nr., Sprachangaben, usw.) im Vordergrund. In der Datenbank *Ergebnisse* beschreiben die Kategorien die Ergebnisse der Forschung, die meldende Organisation, die Rechte am entsprechenden Forschungsergebnis sowie die Form der gewünschten Zusammenarbeit. Die inhaltliche Erschließung erfolgt in beiden Datenbanken mit einem normierten Vokabular, dem sogenannten Themenindexcode.¹⁸² Zusätzliche Angaben können in den Kategorien Facetten-Codes (Datenbank *Veröffentlichungen*) sowie Themendeskriptoren und Themenklassifikation (Datenbank *Ergebnisse*) erfaßt

¹⁷⁹ Bzgl. der Benutzbarkeit der Datenbanken vgl. den Abschnitt zu den „Form Criteria“ in diesem Kapitel.

¹⁸⁰ Die technischen Angaben zu beiden Datenbanken finden sich unter http://dbs.cordis.lu/DE_PUBLI_search.html (Datenbank *Veröffentlichungen*) sowie http://dbs.cordis.lu/DE_RESUI_search.html (Datenbank *Ergebnisse*); beide Stand 1998.

¹⁸¹ Seiteninformationen unter http://dbs.cordis.lu/DE_PUBLI_search.html (Datenbank *Veröffentlichungen*) sowie http://dbs.cordis.lu/DE_RESUI_search.html (Datenbank *Ergebnisse*).

sein.¹⁸³ Nutzer können daher auf eine sehr weit gefaßte Erschließung zurückgreifen, wodurch die Katalogisierung den Bedürfnissen der Zielgruppe angepaßt sein dürfte. Aus den Zielen und Zielgruppen des CORDIS-Systems, ergibt sich der "Geographical Issue" quasi automatisch. Beide Datenbanken beschränken sich auf die Länder der Europäischen Union. Die Datensätze beider Datenbanken sind in englischer Sprache verzeichnet, auch die Abfragesprache der Datenbank ist englisch. In der Datenbank *Veröffentlichungen* sind z.T. die Originaltitel oder die Verfügbarkeit der Dokumente in anderen Sprachen angegeben. Da sich beide Datenbanken an europäische Unternehmen richten und Englisch mittlerweile zum alltäglichen Sprachgebrauch gehört, kann davon ausgegangen werden, daß die Verzeichnung der Datensätze in englischer Sprache den Bedürfnissen der Zielgruppe entspricht.

In einem zweiten Schritt sollen mit Hilfe der **Content Criteria** die Inhalte der beiden CORDIS-Datenbanken beurteilt werden. Dabei sind die Qualität der eigentlichen Daten und die Qualität der Einträge zu unterscheiden. Die Selection criteria fordern dabei im einzelnen sowohl die „Validity“ (Gültigkeit), die „Authority and Reputation“ (Autorität und Reputation) als auch die „Substantiveness“ (Stichhaltigkeit), die „Accuracy“ (Genauigkeit) sowie „Uniqueness“ (Einmaligkeit) und „Comprehensiveness“ (Vollständigkeit) der verzeichneten Informationen. Zusätzlich muß auf die Gestaltung und Organisation (Composition and Organisation) geachtet werden.

Im Bereich der „Validity“ soll u.a. der Wahrheitsgehalt der Informationen geprüft, im Bereich der „Authority and Reputation“ festgestellt werden, inwieweit die in den Datenbanken verzeichneten Informationen gefiltert wurden.

In beiden Datenbanken werden die Einträge vor dem Einspielen geprüft, was am sogenannten „Datum der Qualitätsbewertung“ erkennbar ist. Von der Gültigkeit, der Autorität und Reputation sowie der Stichhaltigkeit, Genauigkeit, Einmaligkeit und Vollständigkeit der verzeichneten Informationen (Qualität und Verlässlichkeit der eigentlichen Daten) könnte demzufolge grundsätzlich ausgegangen werden. Problematisch ist in diesem Zusammenhang jedoch, daß die verwendeten Prüfkriterien nicht offengelegt werden. Verlässliche Aussagen könnten erst dann getroffen werden, wenn bekannt wäre, ob die Angaben der Datenbanken nur bzgl. ihrer Vollständigkeit oder auch inhaltlich geprüft werden.

¹⁸² List of CORDIS Subject Index Classification Codes (1997): <http://dbs.cordis.lu/help/de/su.htm>.

¹⁸³ Diese Angaben finden sich in den Datenbankbeschreibungen unter http://www.cordis.lu/de/src/a_017_de.htm (Veröffentlichungen) und unter http://www.cordis.lu/de/src/a_018_de.htm (Ergebnisse), beide aus 1998.

Im folgenden wird die Qualität der Einträge der beiden CORDIS-Datenbanken bewertet.

In den Einträgen der Datenbanken wurden keine Schreibfehler gefunden, so daß von einer Genauigkeit ausgegangen werden kann. Bezüglich der im Bereich der Stichhaltigkeit geforderten Prüfung von Mehrwerten ("*value-added information*"), ist festzustellen, daß die Einträge in beiden Datenbanken Querverbindungen zu anderen Datenbanken des CORDIS-Systems beeinhalteten können und ihren Nutzern damit die Möglichkeit geben, zusätzliche Informationen, z.B. zu einem EU-Projekt, direkt abzufragen. Derartige Informationen müßten ohne diese Funktionen gesondert abgefragt werden. Die Verfasserin ist deshalb der Ansicht, daß der geforderte Mehrwert der Einträge vorliegt.

Zur Prüfung der Vollständigkeit (**Comprehensiveness**) wurden in beiden Datenbanken thematische Recherchen durchgeführt. Dabei zeigte sich, daß die Datenbanken z.T. sehr ausführliche Abstracts (Datenbank *Veröffentlichungen*) oder Beschreibungen des Forschungsergebnisses enthalten.¹⁸⁴ Unvollständige Einträge wurden im Rahmen der durchgeführten Recherchen nicht ermittelt, obwohl dem Dokumentationsmaterial der Datenbank *Veröffentlichungen* zu entnehmen ist, daß in der Datenbank unvollständige Einträge verzeichnet sein können. Dabei handelt es sich aber nicht um „vergessene Angaben“, sondern um vorläufige Einträge, die erstellt werden, um den Nutzern möglichst frühzeitig Informationen zu den Forschungs- und Entwicklungsprogrammen geben zu können.¹⁸⁵ Beide CORDIS-Datenbanken verzeichnen genau die Informationen, mit denen ein Nutzer, entsprechend den Angaben im Dokumentationsmaterial, rechnen konnte. Insgesamt dürfte die Vollständigkeit der Einträge beider Datenbanken damit vorliegen.

CORDIS ist das einzige Datenbanksystem seiner Art, von einer Einmaligkeit der beiden Datenbanken (Uniqueness) wird daher ausgegangen.

Als letzter Bereich der Content Criteria soll die Gestaltung und Organisation (**Composition and Organisation**) der Informationen beurteilt werden. Die Einträge in beiden Datenbanken sind übersichtlich und logisch aufgebaut. Sie dürften damit den Bedürfnissen ihrer Nutzer nach schneller und gezielter Information entsprechen. In der Datenbank *Ergebnisse* war der Aufbau aller recherchierten Datensätze gleich, womit auch eine Konsistenz der Einträge vorhanden ist. In der Datenbank *Veröffentlichungen* können die einzelnen Datensätze in ihrem Format aufgrund einer

¹⁸⁴ Zur Breite und Tiefe der Indexierung, vgl. den Abschnitt zur „Scope Policy“.

Umstellung der Eintragsform im Jahr 1990 z.T. unterschiedlich sein.¹⁸⁶ Da im Dokumentationsmaterial aber darauf hingewiesen wird, kann die dadurch im Vergleich zur Datenbank *Ergebnisse* entstehende geringere Konsistenz nach Meinung der Verfasserin außer Betracht bleiben. Demzufolge erfüllen beide Datenbanken auch bzgl. der Gestaltung und Organisation die Forderungen der Selection criteria for quality controlled information gateways.

In einem dritten Schritt soll mit Hilfe der **Form Criteria** das System der beiden CORDIS-Datenbanken beurteilt werden. Folglich war auf die Benutzbarkeit (Ease of Navigation) der beiden Datenbanken und darüber hinaus auf das Vorhandensein von Nutzerunterstützung (Provision of User support) sowie eine ästhetische Gestaltung (Aesthetics) der Datenbanken zu achten.¹⁸⁷

Zunächst soll die **Benutzbarkeit** der Datenbanken beurteilt werden. Eine Recherche kann in beiden CORDIS-Datenbanken in freien Suchfeldern oder in Feldern mit Indices erfolgen. Im Bereich der "freien Suche" ist zusätzlich eine Verknüpfung von Suchbegriffen durch Boolesche Operatoren oder eine Phrasensuche möglich. Dies entspricht den üblichen Standardfunktionen einer Datenbank. Die Recherche mit Hilfe der Indices ist allerdings sehr komfortabel. Soll nach bestimmten Begriffen recherchiert werden, ist es ausreichend, den Anfangsbuchstaben des Begriffs einzugeben, um zur entsprechenden Stelle im Index verbunden zu werden. Das Suchwort muß dann nur noch markiert werden. Eine gleichzeitige Recherche nach mehreren Begriffen aus demselben Index ist ebenfalls möglich.

Obwohl Englisch, die eigentliche Abfragesprache darstellt, ist es dennoch möglich die Feldbeschreibung in beiden Datenbanken alternativ auch in französischer, italienischer oder deutscher Sprache anzeigen zu lassen. Dies ist sehr benutzerfreundlich. Die Datenbank *Veröffentlichungen* zeichnet sich außerdem dadurch aus, daß ein gezieltes Einschränken erzielter Suchergebnisse durch die Angabe von Erscheinungsjahren möglich ist. In der Datenbank *Ergebnisse* besteht die Möglichkeit, Suchergebnisse mit Hilfe der sogenannten "landkarten-gestützten Suche" (Map based Search) auf bestimmte Regionen der Europäischen Union, in der Regel Gebiete oder Kreise, zu beschränken. Grundlage dieser kartengestützten Recherche ist die sogenannte "Nuts-Klassifikation" von Eurostat, dem

¹⁸⁵ Vgl. die Datenbankbeschreibung (1998) unter http://www.cordis.lu/de/src/a_017_de.htm.

¹⁸⁶ Vgl. Datenbankbeschreibung zur Datenbank Veröffentlichungen (1998) unter http://www.cordis.lu/de/src/a_017_de.htm.

¹⁸⁷ Bzgl. der außerdem geforderten Verwendung von anerkannten Standards und angemessener Technologie, vgl. den Abschnitt zur „Scope Policy“.

Europäischen Amt für Statistik. Die entsprechenden Gebiete können durch „anklicken“ in einer Europakarte markiert werden. Beide CORDIS-Datenbanken zeichnen sich insgesamt durch eine sehr gute Benutzbarkeit aus.

Die Gestaltung (**Aesthetics**) beider Datenbanken ist nach Meinung der Verfasserin, ebenfalls positiv zu bewerten. Sowohl in der Datenbank als auch in der Datenbank *Veröffentlichungen* trägt die Gestaltung zur Übersichtlichkeit der Suchmasken bei. Diese Übersichtlichkeit wird lediglich in der Datenbank *Veröffentlichungen* dadurch geringfügig beeinträchtigt, daß die Suchmaske mehr als den Umfang einer Bildschirmseite umfaßt.

Der dritte Bereich, der in dieser Kategorie untersucht werden soll, ist die angebotene Nutzerunterstützung (**User support**). Beide Datenbanken bieten ihren Nutzern verschiedene Formen der Unterstützung an. Sowohl Dokumentationsmaterialien in Form von Datenbankbeschreibungen als auch ein Help Desk sind vorhanden. Durch die Datenbankbeschreibung wird dem Nutzer sowohl die Funktion bzw. Bedeutung der einzelnen Felder als auch der thematische und zeitliche Umfang der verzeichneten Daten erläutert. Diese Dokumentationsmaterialien sind neben englischer auch in französischer, italienischer und deutscher Sprache erhältlich, wodurch Nutzern der Datenbanken die Handhabung des Materials wesentlich erleichtert werden kann. Im Rahmen der vorliegenden Arbeit wurden die deutschsprachigen Beschreibungen verwendet, welche allerdings aufgrund von Übersetzungsproblemen z.T. schwer verständlich sind.

Bezüglich der Qualität des Dokumentationsmaterials sollte ein Aspekt aber unbedingt berücksichtigt werden. Festzustellen ist in diesem Zusammenhang, daß die Dokumentationsmaterialien beider Datenbanken nicht offenlegen, nach welchen Kriterien die Auswahl der verzeichneten Informationen erfolgt. Nutzer erfahren z.B. nicht, ob grundsätzlich alle Publikationen zu EG-finanzierten Forschungs- und Entwicklungsprogrammen in der Datenbank *Veröffentlichungen* erfaßt sind, oder nach welchen Merkmalen eine eventuelle Einschränkung erfolgt. Hierin liegt ein wichtiger Qualitätsmangel beider Datenbanken.¹⁸⁸

Eine weitere Form der Nutzerunterstützung bieten beide Datenbanken mit der Funktion "Erläuterung Ihrer Suche", die immer dann angeboten wird, wenn eine Recherche zu keinem Ergebnis führt. Dabei werden die der Recherche zugrunde gelegten Themen oder Begriffe nochmals angezeigt. Nutzer erfahren hierbei

¹⁸⁸ Vgl. Kapitel I.1.1.

allerdings nicht, ob das fehlende Suchergebnis in nicht verzeichneten Dokumenten oder einer falschen Suchanfrage begründet ist. Schulungen oder Training werden für beide Datenbanken nicht angeboten, aber aufgrund der Einfachheit in der Benutzung auch nicht notwendig.

Im Bereich der Nutzerunterstützung positiv hervorzuheben, ist die direkte Erreichbarkeit des Help Desk. In beiden Datenbanken wird der Nutzer über den Menue-Punkt "Kontakte" unmittelbar zur Email-Adresse des Help Desk verbunden, die dann nur noch "angeklickt" werden muß.

Im Bereich der **Process Criteria** soll die Integrität der Information also der verzeichneten Daten (Information Integrity), der Datenbank (Site Integrity) und des Systems (System Integrity) beurteilt werden. Im einzelnen sind die Aktualität und Dauerhaftigkeit der Informationen und Datenbanken sowie die technische Leistungsfähigkeit und Stabilität des Systems zu prüfen. Es ist ebenso zu untersuchen, inwieweit die Verwaltung der Daten und des Systems, bezogen auf die Bedürfnisse der Zielgruppen der Datenbanken, angemessen sind.

Die Prüfung der Aktualität der Informationen und Datenbanken erwies sich in beiden CORDIS-Datenbanken als schwierig. Aufgrund der vorgegebenen Maskenstruktur kann nicht festgestellt werden, zu welchem Zeitpunkt die Datensätze in die Datenbanken eingespielt wurden. Verlässliche Aussagen bezüglich der Aktualität der Informationen sind deshalb nicht möglich. Im Rahmen stichprobenartig durchgeführter thematischer Recherchen konnte allerdings festgestellt werden, daß in der Datenbank *Ergebnisse* Datensätze mit einem „Datum der Aktualisierung“ aus 1996 verzeichnet sind. Dies widerspricht den Angaben der Datenbankbeschreibung, wonach die enthaltenen Daten maximal ein Jahr alt sind.¹⁸⁹ Daraus könnte auf eine mangelnde Aktualität der Daten geschlossen werden. Da die Aktualität der Datenbank insgesamt von einer Aktualität ihrer Daten abhängig ist, können auch bezüglich der Aktualität der Datenbanken bedauerlicherweise keine verlässlichen Angaben gemacht werden.

Da es sich bei den CORDIS-Datenbanken um einen EU-finanzierten Dienst handelt und keinerlei Anzeichen erkennbar waren, welche die Dauerhaftigkeit der CORDIS-Datenbanken in Frage gestellt hätten, muß von einer Dauerhaftigkeit des Systems ausgegangen werden.

Ein letzter Bereich der Process Criteria der in beiden Datenbanken zu untersuchen war, ist die technische Leistungsfähigkeit und die Stabilität des Systems (System

¹⁸⁹ Datenbankbeschreibung (1998) unter http://www.cordis.lu/de/src/a_018_de.htm.

Integrity). Beide Datenbanken waren im gesamten Bearbeitungszeitraums der vorliegenden Arbeit erreichbar und ohne Störungen zu benutzen. Im Vergleich zu anderen Internet-Datenbanken, zeichnen sich beide CORDIS-Datenbanken durch sehr kurze Antwortzeiten aus. Die geforderte System Integrity ist nach Meinung der Verfasserin somit gegeben und insgesamt auch von einer angemessenen Verwaltung der Daten und des Systems auszugehen.

VI. Ergebnisse

Im folgenden sollen die wichtigsten Ergebnisse der vorliegenden Arbeit zusammenfassend dargestellt werden.

Die SIGLE-Datenbank zeichnet sich durch eine sehr große Genauigkeit ihrer Datensätze aus, die sich im Rahmen der durchgeführten Recherchen als praktisch fehlerfrei erwiesen. Ebenfalls zu betonen, ist die Konsistenz der Einträge und der Indexierung insgesamt. Auch die gute Zugänglichkeit der Datenbank muß erwähnt werden.

Beide CORDIS-Datenbanken zeichnen sich gleichermaßen durch eine gute Zugänglichkeit und Benutzbarkeit aus. Die Erschließung der Daten, die durch eine große Anpassung an die Bedürfnisse der Zielgruppe der Datenbanken gekennzeichnet ist, sollte ebenfalls hervorgehoben werden.

Dennoch wurden im Rahmen dieser Arbeit sowohl für die SIGLE-Datenbank als auch für die beiden CORDIS-Datenbanken Qualitätsmängel festgestellt, die unbedingt zu beachten sind und daher nachfolgend nochmals dargestellt werden.

Im Bereich der SIGLE-Datenbank ist zunächst die schlechte Aktualitätsrate zu nennen. Zur Ermittlung aktueller Literatur, insbesondere von Tagungsberichten oder Reports, kann die SIGLE-Datenbank nach Meinung der Verfasserin nicht empfohlen werden. Zudem wurde eine geringe Endnutzerorientierung der Datenbank festgestellt. Festzuhalten ist dabei, daß Endnutzer möglicherweise sogar bei der Nutzung der STN Easy-Version von SIGLE benachteiligt sein können.

Bei den beiden CORDIS-Datenbanken entstehen die Qualitätsmängel dadurch, daß Nutzer bzgl. der verwendeten Auswahl- und Prüfkriterien im unklaren gelassen werden.

Abschließend ist festzustellen, daß sowohl die Ansätze zur Qualitätsbeurteilung der SCOUG als auch die Selection Criteria for Quality Controlled Information gateways ohne Schwierigkeiten angewandt werden konnten. Sie sind zur Qualitätsbeurteilung von Datenbanken uneingeschränkt geeignet. Beiden Modellen ist eine weitere Verbreitung, insbesondere auch in Deutschland, zu wünschen.

VII. Anhang

1. Die Ansätze des Komitees zur Qualitätsbeurteilung von Datenbanken der Finnish Society for Information Services¹⁹⁰

¹⁹⁰ Juntunen et. al. (1995), S. 205-219.

2. Die Ansätze der Southern California Online User Group (SCOUG)¹⁹¹

¹⁹¹ Basch (1990 II), S. 18-23

3. Bibliographic Database Level des CIQM

4. Selection criteria for quality controlled information gateways

5. Literatur- und Quellenverzeichnis

5.1 Literaturverzeichnis

Aitchison, T. M.:

Aspects of quality.

In: Information Services & Use

8 (1988) 2-4 (Miles Conrad Memorial
Lecture), S. 49-61

Armstrong, Chris. J.:

Database and quality : why not try

'what you see is what you get'?

In: Managing Information

1 (1994) 11/12, S. 28-30

Basch, Reva:

An overview of quality and value in
information services.

In: Basch, Reva [Ed.]:

Electronic Information Delivery : Ensuring Quality
and Value. – Aldershot [u. a.] : Gower, 1995, S. 1-10
ISBN 0-566-07567-9

Basch, Reva:

The seven deadly sins of online services.

In: Online 16 (1992) 4, S. 22-25

Basch, Reva:

[1990 II]]

Measuring the quality of data : Report on the
Fourth Annual SCOUG Retreat.

In: Database Searcher 6 (1990) 8, S. 18-23

Collins, Boyd R.:

Beyond Crusing : Reviewing.

In: Library Journal 120 (1996) February 15, S. 122-124

Cung, Soon M. ; Lee, Juhn Young:

Information discovery on the Internet.

In: Encyclopaedia of Library and Information
Science. – New York: Dekker.

Vol. 62 (1998) Supl. 25, S. 146-157

ISBN 0-8247-2062-8

Damerau, Fred:

A Technique for Computer Detection
and Correction of Spelling Errors.

In: Communications of the ACM

7 (1964) 3, S.171-176

Ebbinghouse, Carol:

Information Professionals face the Internet :

The Ninth Annual SCOUG Retreat.

In: Searcher 3 (1995) 9, S. 48-56

Garvin, David A.:

Managing Quality : the strategic
and competitive edge.

New York : The Free Press, 1988

ISBN 0-02-911380-6

Gaus, Wilhelm:

Dokumentation und Ordnungslehre : Theorie
und Praxis des Information Retrieval. – 2. völlig
neu bearb. Aufl. – Berlin [u. a.] : Springer, 1995

ISBN 3-540-58117-0

Gilchrist, Alan:

Quality issues in the information sector: an international
perspective with particular reference to the European scene.

In: Basch, Reva [Ed.]:

Electronic Information Delivery : Ensuring Quality
and Value. – Aldershot [u. a.] : Gower, 1995, S.245-259

ISBN 0-566-07567-9

Granick, Lois:

Assuring the quality of information dissemination :
responsibilities of database producers.

In: Information Services & Use

11 (1991) 1/2, S. 117-136

Haller, Sabine:

Methoden zur Beurteilung von Dienstleistungsqualität.

In: Zeitschrift für betriebswirtschaftliche Forschung

45 (1993), 1, S. 19-40

Jascó, Péter:

Content evaluation of databases.

In: Annual Review of Information

Science and Technology (ARIST)

32 (1997), S. 231-267

Jascó, Péter:

[1993 I]

Searching for skeletons in the database cupboard Part I : Errors of omission.

In: Database 16 (1993) 1 S. 38-49

Jascó, Péter:

[1993 III]

Searching for skeletons in the database cupboard. Part II : Errors of commission.

In: Database 16 (1993) 2, S. 30-36.

Jascó, Péter:

[1993 II]

A Proposal for database „Nutrition and Ingredient“ Labeling.

In: Database 16 (1993)1, S. 7-9

Jascó, Péter:

What is in a(n) (Up)date?

Currency test searching of databases.

In: Database 15 (1992) 3, S. 28-33

Johannsen, Carl Gustav:

Quality Management principles and methods in Library and Information Science Theory and Practice.

In: Encyclopaedia of Library and Information Science. - New York: Dekker.

Vol. 56, Suppl. 19 (1995), S. 318-352

ISBN 0-8247-2056-3

Juntunen, Ritva; Mickos, Elisabet ;

Jalkanen, Tuulikki:

Evaluating the quality of Finnish databases.

In: Basch, Reva [Ed.]:

Electronic Information Delivery : Ensuring Quality and value. – Aldershot [u. a.] : Gower, 1995, S. 205-219

ISBN 0-566-07567-9

Marchand, Donald:

Managing Information Quality.

In: Wormel, Irene [Ed.]:

Information Quality : definitions and dimensions. -London [u.a.] :

Taylor Graham, 1990, S. 7-17

ISBN 0-947568-43-3

Medawar, Katia:

Database Quality: a literature review
of the past and a plan for the future.

In: Program 29 (1995) 3, S. 257-272

Mintz, Anne P.:

Quality Control and the Zen of
Database Production.

In: Online 14 (1990) 6, S. 15-23

Mitton, Roger:

Spelling Checkers, Spelling Correctors
and the Misspelling of Poor Spellers.

In: Information Processing & Management
23 (1987) 5, S. 495-505.

**Neubauer, Wolfram [Hrsg.] ; Deutsche
Gesellschaft für Dokumentation [Veranst.]:**

Online und darüber hinaus...Tendenzen der
Informationsvermittlung (17. Online-Tagung
der DGD Frankfurt am Main 16. bis 18. Mai
1995). Frankfurt am Main : DGD, 1995. -
Originalbeitrag

Norpoth, Susanne ; Nüßlein, Désirée:

Datenbanken in der Marktforschung aus Sicht
eines Nutzers, S. 201-216

Norm EN ISO 8402 : 1995.

Qualitätsmanagement : Begriffe

O'Neill, Edward T. ; Viziné-Goetz, Diane:

Quality control in online databases.

In: Annual Review of Information Science and
Technology (ARIST) 23 (1988), S. 125-156

**Parasuraman, A. ; Zeithaml, Valerie A. ; Berry,
Leonard L.:**

SERVQUAL : a multiple-item Scale for measuring
consumer perceptions of Service Quality.

In: Journal of Retailing 64 (1988) 1, S. 5-37

Pepels, Werner:

Qualitätscontrolling bei Dienstleistungen.

München : Vahlen, 1996

ISBN 3-8006-2133-9

**Raitt, David I. [Ed.] ; Learned Information
(Europe) Ltd. [Veranst.]:**

Online Information 90 (14th International
Online Information Meeting London 11 –
13 December 1990). Oxford [u.a.] :
Learned Information, 1990. –Originalbeitrag

Basch, Reva:

[1990 I]

Finding our voice : the user 'wish list'
and the building of a searcher community,
S. 251-259

**Raitt, David I. [Ed.] ; Learned Information
(Europe) Ltd. [Veranst.]:**

Online Information 91 (15th International Online
Information Meeting London 10 – 12 December 1991).
Oxford [u.a.] : Learned Information, 1991. -
Originalbeitrag

[Juntunen et. al. (1991)]

Quality requirements for databases : project for
evaluating Finnish databases / R. Juntunen..., S. 351-359

**Raitt, David I. [Ed.] , Learned Information
(Europe) Ltd. [Veranst.]:**

Online Information 95 (19th International Online
Information Meeting London 5 –7 December
1995). Oxford [u.a.] : Learned Information, 1995. –
Originalbeitrag

Armstrong, Chris J.:

Do we really care about quality?, S. 49-59

**Raitt, David I.[Ed.] ; Learned Information
Europe Ltd. [Veranst.]:**

Online Information 96 (20th International Online
Information Meeting 3 – 5 December 1996).
Oxford [u.a.] : Learned Information, 1996. –
Originalbeiträge

Cooke, Alison ; McNab, Alison,

Anagnostelis, Betsy:

The good, the bad and the ugly :

Internet review sites, S. 33-40

Ochsner, Heinz ; Thomas K.:

The Internet : the demise of
traditional online, S. 477-482

**Schwuchow, Werner [Hrsg.] ; Kommission
Wirtschaftlichkeit der Information und Doku-
mentation in der Deutschen Gesellschaft
für Dokumentation [Veranst.] ; Gesellschaft für
Informatik e.V. [Veranst.] ; International Federation
for Information and Documentation [Veranst.]:**

Qualität von Informationsdiensten (7. internationale
Fachkonferenz der Kommission Wirtschaftlichkeit der
Information und Dokumentation KWID in der Deutschen
Gesellschaft für Dokumentation in Zusammenarbeit
mit der Gesellschaft für Informatik e.V. GI und der
International Federation for Information and Docu-
mentation FID Garmisch-Partenkirchen

2. – 4. Mai 1993). Frankfurt : DGD, 1993. –
Originalbeiträge

Keiser, Barbie E.:

Quality Management in Information Services : a North
American Perspective, S. 40-53.

Stroethmann, Karl A.:

Qualität von Informationsdiensten als
strategische Managementaufgabe, S. 10-22

Su, Louise, T.:

Evaluation Measures for interactive
Information Retrieval.

In: Information Processing & Management
28 (1992) 4, S. 503-516

Tenopir, Carol:

Database Quality revisited.

In: Library Journal 115 (1990) Okt. 1, S. 64-67

Wilson, T. D.:

EQUIP: a European survey of quality
criteria for the evaluation of databases.

In: Journal of Information Science
24 (1998) 4, S. 345-357

5.2 Quellenverzeichnis

Alexander, Jan / Tate, Marsha Ann:

Web ressource evaluation.
<http://www.widener.edu/~withers/webeval.htm>
(Einsichtnahme am 29.07.1999)

→The→ Argus Clearinghouse
<http://www.clearinghouse.net>
(Einsichtnahme am 23.09.1999)

Armstrong, Chris J.:
Quality on the Internet.
In: db-Qual 2 (1997)1
http://www.i-a-l.co.uk/CIQM/db_qual/2_1_4.htm
Stand 13. Januar 1999
(Einsichtnahme am 28.07.1999)

Armstrong, Chris J.:
The Quality of publicly-available
databases: WYSIWYG or what?
Stand 13. Januar 1999
<http://www.i-a-l.co.uk/ciqm/BULISC.htm>
(Einsichtnahme am 07.05.1999)

Bibliographic Database Label
ISI SCISearch Online.
<http://www.la-hq.org.uk/liaison/ciqm/isisso.html>
Stand Oktober 1996
(Einsichtnahme am 23.09.1999)

**Botte, A. [Hrsg.] ; Gesellschaft
Information Bildung [Veranst.]:**
Schritte zur Qualitätsverbesserung von
Bildungsinformationssystemen (5. GIB-
Fachtagung, 1. und 2. Oktober 1997
in Halle). Berlin : GBI, 1998. – Originalbeitrag

Hobohm, Hans-Christoph:
Qualität und Kundenzufriedenheit als Management-
kriterien für Informationseinrichtungen
Preprint <http://www.fh-potsdam.de/~hobohm/gib.pdf>

Ciolek, Matthew T.:
Information Quality - Catalogue of Potent Truisms.
<http://www.ciolek.com/WWWWWVLPages/QtyPages/Qty/Truisms.html>
Stand 10. November 1997
(Einsichtnahme am 30.08.1999)

Grassian, Esther:

Thinking critically about World Wide Web Ressources.

<http://www.library/ucla.edu/libraries/college/instruct/web/critical.htm>

Stand 01. April 1999

(Einsichtnahme am 29.07.1999)

Hiom, Debra:

The Social Science Information Gateway :

Putting Theory into Practice.

In: Information Research 4 (1998)1

<http://www.shef.ac.uk/~is/publications/infres/paper48.html>

Stand 10. Juni 1998

(Einsichtnahme am 21.07.1999)

Hofmann, Paul ; Worsfold, Emma:

Specifications for resource description methods

Part 2: Selection Criteria for Quality Controlled Information Gateways.

<http://www.ukoln.ac.uk/metadata/desire/quality/toc.html>

Stand 02. April 1998

(Einsichtnahme am 21.07.1999)

Infofilter-Projekt

<http://www.usc.edu/user/help/flick/Infofilter>

(Einsichtnahme am 23.09.1999)

Information Market Observatory (IMO) [Hrsg.]:

? The? Quality of Electronic Information Products and Services:

Luxemburg, September 1995

(IMO Working Paper ; 95/4)

<http://www.echo.lu/impact/imo/9504.html>

(Einsichtnahme am 05.07.1999)

Iser, Uschi <ui@fiz-karlsruhe.de>. 14.09.1999

Re: „Qualitätssicherung SIGLE, Ihre Anfrage“

Persönliche E-Mail (14.09.1999)

Kirk, Elizabeth:

Evaluating information found on the Internet.

<http://milton.mse.jhu.edu:8001/research/educaton/net.html>

Stand 06. Februar 1999

(Einsichtnahme am 05.08.1999)

Payer, Margarete <1942- >:

Wie kann man die Qualität von Internetressourcen für den wissenschaftlichen Bereich beurteilen? : Hinterfragung ausgewählter Vorschläge ; Vortrag am 11. Juni 1997, HBI Stuttgart. - Fassung vom 09. Juni 1997.

<http://machno.hbi-stuttgart.de/?payer/infoq.html>.

(Einsichtnahme am 05.08.1999)

Rettig, James:

Beyond "Cool" : Analog Models for Reviewing Digital Resources.

in: Online, September 1996

<http://www.onlineinc.com/onlinemag/septOL/rettig9.html>

Stand September 1996

(Einsichtnahme am 26.08.1999)

Swindells, Norman:

Managing the Quality of Information Products.

Stand September 1995

<http://www.echo.lu/qa/en/qa2.html>

(Einsichtnahme am 13.07.1999)

The WebCrawler Guide

<http://webcrawler.com/Help/Guide.html>

(Einsichtnahme am 30.08.1999)

Yahoo! Internet Life

<http://www.zdnet.com/yil/content/depts/stargde.html>

(Einsichtnahme am 30.08.1999)

5.3 Verzeichnis der verwendeten Dokumentations- und Hilfsmaterialien**5.3.1 Materialien zur CORDIS-Datenbank****CORDIS-Homepage**

<http://www.cordis.lu>

Stand 1999

(Einsichtnahme 16.08.1999)

CORDIS : Ihr Schlüssel zur Innovation
[CORDIS-Einführung]
http://www.cordis.lu/de/src/s_016_de.htm
Stand 18. Januar 1999
(Einsichtnahme am 28.07.1999)

Datenbank FTE-Ergebnisse
http://www.cordis.lu/de/src/a_018_de.htm
Stand 05. Oktober 1998
(Einsichtnahme am 30.08.1999)

Datenbank FTE-Veröffentlichungen
http://www.cordis.lu/de/src/a_017_de.htm
Stand 05. Oktober 1998
(Einsichtnahme am 30.08.1999)

**List of CORDIS Subject Index
Classification Codes**
<http://dbs.cordis.lu/help/de/su.htm>
Stand 1997
(Einsichtnahme am 23.09.1999)

5.3.2 Materialien zur SIGLE-Datenbank

Customer Service
FIZ/STN Global Hotline
<http://www.fiz-karlsruhe.de/stn.html>
Stand 25. Juni 1998
(Einsichtnahme am 23.09.1999)

**Detailed Description of PC Hardware
Components for Online Searching**
<http://www.fiz-karlsruhe.de/fiz/faq/hardware.html>
Stand 27. März 1997
(Einsichtnahme am 07.09.1999)

**How to connect to STN International
in Europe**
<http://www.fiz-karlsruhe.de/stn.html>
Stand Juni 1998
(Einsichtnahme am 23.09.1999)

SIGLE Description
<http://www.FIZ-Karlsruhe.DE/stn/Databases/sigle.html>

Stand 06.08.1999
(Einsichtnahme am 12.08.1999)

SIGLE STN Database Summary Sheet

<http://info.cas.org/ONLINE/DBSS/sigless.html>

Stand Juli 1999
(Einsichtnahme am 16.08.1999)

STN Easy Hilfe

[STN Easy-Dokumentation]

<http://stneasy.fiz-karlsruhe.de/html/deutsch/help/.html#about>

(Einsichtnahme am 23.09.1999)

STN Newsletter

April 1999

STN Terms and Conditions

<http://stnweb.fiz-karlsruhe.de>

Stand 1999
(Einsichtnahme am 07.09.1999)

5. Literatur- und Quellenverzeichnis

5.1 Literaturverzeichnis

Aitchison, T. M.:

Aspects of quality.

In: Information Services & Use
8 (1988) 2-4 (Miles Conrad Memorial
Lecture), S. 49-61

Armstrong, Chris. J.:

Database and quality : why not try
'what you see is what you get'?
In: Managing Information
1 (1994) 11/12, S. 28-30

Basch, Reva:

**An overview of quality and value in
information services.**

In: Basch, Reva [Ed.]:
Electronic Information Delivery : Ensuring Quality
and Value. – Aldershot [u. a.] : Gower, 1995, S. 1-10

Basch, Reva:

The seven deadly sins of online services.
In: Online 16 (1992) 4, S. 22-25

Basch, Reva:

Measuring the quality of data : Report on the
Fourth Annual SCOUG Retreat.
In: Database Searcher 6 (1990) 8, S. 18-23

Collins, Boyd R.:

Beyond Crusing : Reviewing.
In: Library Journal 120 (1996) February 15, S. 122-124

Cung, Soon M. ; Lee, Juhn Young:

Information discovery on the Internet.

**In: Encyclopeadia of Library and Information
Science. – New York: Dekker.**

0-8247-2062-8

Vol. 62 (1998) Supl. 25, S. 146-157

Damerau, Fred:

A Technique for Computer Detection
and Correction of Spelling Errors.
In: Communications of the ACM
7 (1964) 3, S.171-176

Ebbinghouse, Carol:

Information Professionals face the Internet :
The Ninth Annual SCOUG Retreat.
In: Searcher 3 (1995) 9, S. 48-56

Garvin, David A.:

Managing Quality : the strategic
and competitive edge.
New York : The Free Press, 1988
ISBN 0-02-911380-6

Gaus, Wilhelm:

Dokumentation und Ordnungslehre : Theorie

und Praxis des Information Retrieval. – 2. völlig
neu bearb. Aufl. – Berlin [u. a.] : Springer, 1995
ISBN 3-540-58117-0

Gilchrist, Alan:

Quality issues in the information sector: an international

perspective with particular reference to the European scene.
In: Basch, Reva [Ed.]:
Electronic Information Delivery : Ensuring Quality
and Value. – Aldershot [u. a.] : Gower, 1995, S.245-259

Granick, Lois:

Assuring the quality of information dissemination :
responsibilities of database producers.
In: Information Services & Use
11 (1991) 1/2, S. 117-136

Haller, Sabine:

Methoden zur Beurteilung von Dienstleistungsqualität.

In: Zeitschrift für betriebswirtschaftliche Forschung
45 (1993), 1, S. 19-40

Jascó, Péter:

Content evaluation of databases.
In: Annual Review of Information
Science and Technology (ARIST)
32 (1997), S. 231-267

Jascó, Péter:

Searching for skeletons in the database cupboard Part I : Errors of omission.

in: Database 16 (1993) 1 S. 38-49

Jascó, Péter:

Searching for skeletons in the database cupboard. Part II : Errors of commission.

in: Database 16 (1993) 2, S. 30-36.

Jascó, Péter:

A Proposal for database „Nutrition and Ingredient“ Labeling.

In: Database 16 (1993)1, S. 7-9

Jascó, Péter:

What is in a(n) (Up)date?

Currency test searching of databases.

In: Database 15 (1992) 3, S. 28-33

Johannsen, Carl Gustav:

Quality Management principles and methods in Library and Information Science Theory and Practice.

In: Encyclopedia of Library and Information Science. -

New York: Dekker.

ISBN 0-8247-2056-3

Vol. 56, Suppl. 19 (1995), S. 318-352

Juntunen, Ritva; Mickos, Elisabet ;

Jalkanen, Tuulikki:

Evaluating the quality of Finnish databases.

In: Basch, Reva [Ed.]:

Electronic Information Delivery : Ensuring Quality and value. – Aldershot [u. a.] : Gower, 1995, S. 205-219

Marchand, Donald:

Managing Information Quality.

In: Wormel, Irene [Ed.]:

Information Quality : definitions and dimensions. -London [u.a.] :

Taylor Graham, 1990, S. 7-17

Medawar, Katia:

Database Quality: a literature review
of the past and a plan for the future.
In: Program 29 (1995) 3, S. 257-272

Mintz, Anne P.:

Quality Control and the Zen of
Database Production.
In: Online 14 (1990) 6, S. 15-23

Mitton, Roger:

Spelling Checkers, Spelling Correctors
and the Misspelling of Poor Spellers.
In: Information Processing & Management
23 (1987) 5, S. 495-505.

Neubauer, Wolfram [Hrsg.]:

Online und darüber hinaus...Tendenzen der
Informationsvermittlung (17. Online-Tagung
der DGD Frankfurt am Main 16. bis 18. Mai
1995). Frankfurt am Main : DGD, 1995. -
Originalbeitrag
Norpoth, Susanne ; Nüßlein, Désirée:
Datenbanken in der Marktforschung aus Sicht
eines Nutzers, S. 201-216

Norm EN ISO 8402 : 1995.

Qualitätsmanagement : Begriffe

O'Neill, Edward T. ; Vizin-Goetz, Diane:

Quality control in online databases.
In: Annual Review of Information Science and
Technology (ARIST) 23 (1988), S. 125-156

Online Information 90 : 14th International
Online Information Meeting ; Proceedings, London
11 – 13 December 1990 / ed. by David I. Raitt. –
Oxford [u.a.] : Learned Information, 1990. –
Originalbeitrag

Basch, Reva:

Finding our voice : the user 'wish list' and
the building of a searcher community, S. 251-259

Online Information 91 : 15th International Online
Information Meeting ; Proceedings ; London
10 – 12 December 1991 / Org. by Learned Information
(Europe) Ltd. Ed. by David I. Raitt. –
Oxford [u.a.], 1991, - Originalbeitrag

Quality requirements for databases : project for
evaluating Finnish databases / R. Juntunen..., S. 351-359

Online Information 95 : proceedings ; London
5 – 7 December 1995 / 19th International Online
Information Meeting. Org. by Learned Information
(Europe) Ltd. Ed. by David I. Raitt... – Oxford [u.a.],
1995. – Originalbeitrag
Armstrong, Chris J.:
Do we really care about quality?, S. 49-59

Online Information 96 : proceedings ; London
3 – 5 December 1996 / 20th International Online
Information Meeting. Org. by Learned Information
(Europe) Ltd. Ed. by David I. Raitt... – Oxford [u.a.],
1996. - Originalbeitrag
Cooke, Alison ; McNab, Alison,
Anagnostelis, Betsy:
The good, the bad and the ugly :
Internet review sites, S. 33-40

**Parasuraman, A. ; Zeithaml, Valerie A. ; Berry,
Leonard L.:**
SERVQUAL : a multiple-item Scale for measuring
consumer perceptions of Service Quality.
In: Journal of Retailing 64 (1988) 1, S. 5-37

Pepels, Werner:
Qualitätscontrolling bei Dienstleistungen.
München : Vahlen, 1996
ISBN 3-8006-2133-9

Schwuchow, Werner [Hrsg.]:
Qualität von Informationsdiensten : 7. internationale
Fachkonferenz der Kommission Wirtschaftlichkeit der
Information und Dokumentation KWID in der Deutschen
Gesellschaft für Dokumentation in Zusammenarbeit
mit der Gesellschaft für Informatik e.V. GI und der International Federation for
Information and Documentation FID ; Garmisch-Partenkirchen, 2. – 4. Mai 1993. -
Frankfurt : DGD, 1993. –
Originalbeiträge

Keiser, Barbie E.:

Quality Management in Information Services : a North

American Perspective, S. 40-53.

Stroethmann, Karl A.:

Qualität von Informationsdiensten als

strategische Managementaufgabe, S. 10-22

Su, Louise, T.:

Evaluation Measures for interactive
Information Retrieval.

In: Information Processing & Management
28 (1992) 4, S. 503-516

Tenopir, Carol:

Database Quality revisited.

in: Library Journal 115 (1990) Okt. 1, S. 64-67

Wilson, T. D.:

EQUIP: a European survey of quality

criteria for the evaluation of databases.

In: Journal of Information Science
24 (1998) 4, S. 345-357

5.2 Quellenverzeichnis

Alexander, Jan / Tate, Marsha Ann:

Web ressource evaluation.

<http://www.widener.edu/~withers/webeval.htm>

(Einsichtnahme am 29.07.1999)

→The→ Argus Clearinghouse

<http://www.clearinghouse.net>

(Einsichtnahme am 23.09.1999)

Armstrong, Chris J.:

Quality on the Internet.

In: db-Qual 2 (1997)1

http://www.i-a-l.co.uk/CIQM/db_qual/2_1_4.htm

Stand Januar 1999

(Einsichtnahme am 28.07.1999)

Armstrong, Chris J.:

The Quality of publicly-available
databases: WYSIWYG or what?
Stand 13. Januar 1999
<http://www.i-a-l.co.uk/ciqm/BULISC.htm>
(Einsichtnahme am 07.05.1999)

Bibliographic Database Label

ISI SCISearch Online.
[http://www.la-hq.org.uk/liaison/
ciqm/isisso.html](http://www.la-hq.org.uk/liaison/ciqm/isisso.html)
Stand Oktober 1996
(Einsichtnahme am 23.09.1999)

Ciolek, Matthew T.:

Information Quality - Catalogue of Potent Truisms.
[http://www.ciolek.com/WWWVLPages/
QtlyPages/Qtly/Truisms.html](http://www.ciolek.com/WWWVLPages/QtlyPages/Qtly/Truisms.html)
Stand 10. November 1997
(Einsichtnahme am 30.08.1999)

Grassian, Esther:

Thinking critically about World Wide Web Ressources.
[http://www.library/ucla.edu/libraries/college/
instruct/web/critical.htm](http://www.library/ucla.edu/libraries/college/instruct/web/critical.htm)
Stand April 1999
(Einsichtnahme am 29.07.1999)

Hiom, Debra:

The Social Science Information Gateway :
Putting Theory into Practice.
In: Information Research 4 (1998)1
[http://www.shef.ac.uk/~is/publications/infres/
paper48.html](http://www.shef.ac.uk/~is/publications/infres/paper48.html)
Stand Juni 1998
(Einsichtnahme am 21.07.1999)

Hobohm, Hans-Christoph:

Qualität und Kundenzufriedenheit als Management-
kriterien für Informationseinrichtungen.
In: Schritte zur Qualitätsverbesserung von Bildungs-
informationssystemen. Proceedings der 5.GIB-

Fachtagung, 1. und 2. Oktober 1997 in Halle, hrsg. von
A. Botte, D. Rush-Feja, und R. Theers. -
Berlin : GBI, 1998, S. 40-59
Preprint <http://www.fh-potsdam.de/~hobohm/gib.pdf>

Hofmann, Paul ; Worsfold, Emma:

Specifications for resource description methods
Part 2: Selection Criteria for Quality Controlled
Information Gateways.
[http://www.ukoln.ac.uk/metadata/desire/
quality/toc.html](http://www.ukoln.ac.uk/metadata/desire/quality/toc.html)
Stand April 1998
(Einsichtnahme am 21.07.1999)

Infofilter-Projekt

<http://www.usc.edu/user/help/flick/Infofilter>
(Einsichtnahme am 23.09.1999)

Information Market Observatory (IMO) [Hrsg.]:

? The? Quality of Electronic Information Products
and Services:
Luxemburg, September 1995
(IMO Working Paper ; 95/4)
<http://www.echo.lu/impact/imo/9504.html>
(Einsichtnahme am 05.07.1999)

Iser, Uschi <ui@fiz-karlsruhe.de>. 14.09.1999
Re: „Qualitätssicherung SIGLE, Ihre Anfrage“
Persönliche E-Mail (14.09.1999)

Kirk, Elizabeth:

Evaluating information found on the Internet.
<http://milton.mse.jhu.edu:8001/research/educaton/net.html>
Stand 06.02.1999
(Einsichtnahme am 05.08.1999)

Payer, Margarete <1942- >:

Wie kann man die Qualität von Internetressourcen für
den wissenschaftlichen Bereich beurteilen? : Hinterfragung
ausgewählter Vorschläge ; Vortrag am 11. Juni 1997, HBI
Stuttgart. - Fassung vom 09. Juni 1997.
<http://machno.hbi-stuttgart.de/?payer/infoq.html>.
(Einsichtnahme am 05.08.1999)

Rettig, James:

Beyond "Cool" : Analog Models for Reviewing
Digital Resources.

in: Online, September 1996

<http://www.onlineinc.com/onlinemag/septOL/rettig9.html>

Stand 1996

(Einsichtnahme am 26.08.1999)

Swindells, Norman:**Managing the Quality of Information Products.**

Stand September 1995

<http://www.echo.lu/qa/en/qa2.html>

(Einsichtnahme am 13.07.1999)

The WebCrawler Guide

<http://webcrawler.com/Help/Guide.html>

(Einsichtnahme am 30.08.1999)

Yahoo! Internet Life

<http://www.zdnet.com/yil/content/depts/stargde.html>

(Einsichtnahme am 30.08.1999)

5.3 Verzeichnis der verwendeten Dokumenta tions- und Hilfsmaterialien

5.3.1 Materialien zur CORDIS-Datenbank

CORDIS-Homepage

<http://www.cordis.lu>

Stand 1999

(Einsichtnahme 16.08.1999)

CORDIS : Ihr Schlüssel zur Innovation

[CORDIS-Einführung]

http://www.cordis.lu/de/src/s_016_de.htm

Stand Januar 1999

(Einsichtnahme am 28.07.1999)

Datenbank FTE-Ergebnisse

http://www.cordis.lu/de/src/a_018_de.htm

Stand Oktober 1998

(Einsichtnahme am 30.08.1999)

Datenbank FTE-Veröffentlichungen

http://www.cordis.lu/de/src/a_017_de.htm

Stand Oktober 1998

(Einsichtnahme am 30.08.199)

List of CORDIS Subject Index

Classification Codes

<http://dbs.cordis.lu/help/de/su.htm>

Stand 1997

(Einsichtnahme am 23.09.1999)

5.3.2 Materialien zur SIGLE-Datenbank

Customer Service

FIZ/STN Global Hotline

<http://www.fiz-karlsruhe.de/stn.html>

Stand Juni 1998

(Einsichtnahme am 23.09.1999)

Detailed Description of PC Hardware

Components for Online Searching

<http://www.fiz-karlsruhe.de/fiz/faq/hardware.html>

Stand März 1997

(Einsichtnahme am 07.09.1999)

How to connect to STN International

in Europe

<http://www.fiz-karlsruhe.de/stn.html>

Stand Juni 1998

(Einsichtnahme am 23.09.1999)

SIGLE Description

<http://www.FIZ-Karlsruhe.DE/stn/Databases/sigle.html>

Stand 06.08.1999

(Einsichtnahme am 12.08.1999)

SIGLE STN Database Summary Sheet

<http://info.cas.org/ONLINE/DBSS/sigless.html>

Stand Juli 1999

(Einsichtnahme am 16.08.1999)

STN Easy Hilfe

[STN Easy-Dokumentation]

<http://stneasy.fiz-karlsruhe.de/html/deutsch/help/.html#about>

(Einsichtnahme am 23.09.1999)

STN Newsletter

April 1999

STN Terms and Conditions

<http://stnweb.fiz-karlsruhe.de>

Stand 1999

(Einsichtnahme am 07.09.1999)

Erklärung

Hiermit erkläre ich, daß ich die vorliegende Diplomarbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich genannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Hünfelden, 07.10.1999